CHAPTER - 2

Review of Literature
**REVIEW OF LITERATURE**

In today’s society, number of professions with high vocal demands is growing rapidly. The amount of vocal demand imposed on these professionals varies across professions. A well-functioning voice is an essential tool for occupational voice users. According to Titze et al., (1997), the concept of a professional voice user has extended from singers and actors to other occupations such as teachers, receptionists, lawyers, broadcasters, priests, telemarketers and counsellors who find it virtually impossible to interact with their clients, students or audiences without a well-functioning and enduring voice. According to Sataloff (1991) professional voice users’ lifestyle and work expose them to many dangers that may jeopardize their most valuable instrument of expression. Their voice problems can be labelled as “occupational voice disorders” as the symptoms they suffer from are likely to be caused by exposure at work. Voice problems in these professionals adversely affect their career or reduce profit for the employer, creating negative effects on their occupational performance and society (Rantala et al., 2002). It’s strange that most of these professional voice users are unaware of their vocal limitations and fail to identify early signs of vocal attrition. They go on and remain professionally productive for a longer period of time with existing vocal misuse or abusive behaviours and seek medical help after long period of onset of the problem (Mitchell, 1996). These professional voice users provide exciting challenges and special responsibilities for physicians, Speech Language Pathologists and other health care professionals, and their need for expert care has inspired new interest in understanding the function and dysfunction of the human voice (Sataloff, 1991).
The definition of a voice disorder in an occupational context depends on the demands set upon the voice. For some professions, an average vocal capacity is sufficient. There are, however, some professionals, like teachers, singers, dramatists and telemarketers who require sustained and extensive use of voice, above and beyond the demands of everyday speaking. As a consequence, these professionals require high degree of vocal capacity. The literature on voice disorders has proposed a variety of definitions of what should be considered as a voice disorder and in general lacks an operational definition of voice problem or disorder (Verdolini & Ramig, 2001). According to ICIDH-2 (2001) (International Classification of Impairments, Disability and Handicaps cited in Lehto, Alku, Backstrom, & Vilkman, 2005), the impact of voice problem depends on a person’s reports of sufficient concern and functional disruption in voice, and his/her opinion should not be ignored. Hence, voice problems should be considered as a multidimensional phenomenon, and interpreted as a complex of self-reported symptoms and clinically observed signs. Supporting this, Laguaite (1976) reported that, there exist high discrepancies in the prevalence of voice problems when judgments are based on clinicians (7%), to that of self-reported symptoms of the subjects (15%). This indicates that, self-perceived voice problems do not necessarily correlate with degree of voice quality impairment as measured acoustically and or identified perceptually. As explained by Lehto et al., (2005), voice disorders always imply voice symptoms experienced as problems, but experiencing symptoms does not always mean that a person has a voice disorder. Hence, in case of occupational voice users, prevalence of voice problems will be significantly underestimated if it is considered solely on the basis of clinical measures. This is relevant at the first signs of voice loading causing vocal fatigue which may emerge as a voice symptom but not showing significant changes in the laryngoscopic, acoustic or perceptual measures.
Thus, in the case of occupational voice users, their own concern about their voice symptoms and related reduction in physical, social, emotional and/or professional well-being should be taken into account while estimating the prevalence of voice problems in order to sustain their occupational capacity (Verdolini & Ramig, 2001). Thus, a self-reported questionnaire study seems to be the most common and appropriate way to investigate the prevalence of occupational voice problems (Herrington-Hall, Lee, Stemple, Niemi & Miller-McHone, 1988; Fritzell, 1996).

The advancement of new speech technology applications in various professional and social contexts has increased the use of voice in professional settings. Lehto, Laaksonen, Vilkman and Alku (2006) commented that telephone services are an example of expanding modern-day speech-related professional contexts with telephone operators constituting a special group of employees whose working ability depends exclusively on their voices. Further, they stated that, on the telephone, the speaker must rely solely on his/her voice without support from body language or written communication. However, telephone customer service advisors get short term recovery while attending to the calls as they are involved in dialogue than monologue like teachers and this form of communication should help them to preserve normal voice function. Literature quotes of higher prevalence of voice problems among employees working in telephone services (Jones et al., 2002; Taylor & Oates, 2004) which reflects that, post loading measurements after short term vocal rest did not return to the pre-loading level and the assumption that a short break is not sufficient for the recovery in voice professionals (Vintturi, Alku, Lauri, Sala, Sihvo & Vilkman, 2001).
2.1 Call centers and work characteristics

Computer based customer service work or call center work is one of the most rapidly growing occupations in the world (Norman, Nilsson, Hagberg, Wigaeus, Tornqvist & Toomingas, 2004). Call Centers are organizations or departments that are specifically dedicated to contacting clients and customers. They have their origin in the USA, where they started in 1908 when it became possible to use the telephone to sell advertisements. In the beginning of the 1960s, Ford Motor Company started to search for possible buyers of their cars by making 20,000,000 phone calls to the consumers (Norman, 2005). Call centers are a unique working environment because of a combination of intensive periods of working with both a telephone and computer. That is, the call center workers have to listen, watch and talk, all at the same time, without a break.

Call center workers can handle either the incoming and/or outgoing telephone calls. Typical services with outgoing calls are advertising campaigns, market research and selling by telephone, while the activities with incoming calls are; customer services, giving information, taking orders and providing help desk, functions. In the last few years, operators have also started to handle e-mail, fax and SMS (short message service) (Norman, 2005). It is estimated that more than 100,000 call centers now exist worldwide. As reported by Paul and Huws (2002), around 1% of work force in Europe works in a call centre, with the greatest call centre employment rates found in the United Kingdom and Ireland. Their proportion in the USA is assumed to be around 3% (Titze et al., 1997).
Call center work stations are usually open plan offices with movable partitions. The operators sit in front of the computer most of the day, with both physically and mentally monotonous and repetitive work. Call center workers often spend 90% of their working time on the telephone and in front of the computer. Thus, computer technology has become a critical component of workplace management in call centers (Batt, 1999) and is used to monitor the speed of work, regulate the level of downtime when the operator is not available to take calls, and assess the quality of the interaction between the service provider and the customer.

Norman (2005) reported that depending on the type of call center, the employees are provided ‘firm-specific human capital’ knowledge. That is, the call handlers manage the boundary between the firm and the customer, and shape the customers buying behaviour. In order to persuade customers to buy a firm’s products and services, employees need a clear understanding of specific product features, service agreements, pricing, packaging, promotions for particular customer segments and legal regulations. He also mentioned their need of customer-specific knowledge regarding the demand characteristics of particular individuals or segments and how to use that knowledge to negotiate customized offerings. In addition, employees also require specific knowledge of the structure and content of the firm’s information systems, the work flow from point of sales to delivery, and how the company’s processing capabilities affect each customer and product offering. However, there is little flexibility in CCOs interaction with the customers as they are required to follow tightly scripted dialogues and conform to highly detailed instructions (Wharton, 1993). Even if the call center operator realizes that the conversation is confusing the customers, the system does not allow the scripts to be changed to meet the needs of the customer (Hannif & Lamm,
The call centers set targets for their call handlers. Targets are usually set in terms of number of calls to be handled by each agent per shift. Call center employees are expected to answer a pre-determined number of calls during their shift. Successful handling of the allotted number of calls constitutes a portion of the performance evaluation of the worker (Bayeh & Smith 1999; Smith & Bayeh, 2003). All these emphasize the strict and regulated work schedule of a CCO and the vocal demands that accompany.

The repetitiveness of tasks and the simultaneous use of computers and telephones represent a number of health related hazards in call center workers such as; back/postural problems, repetitive strain injury, voice loss, acoustic shock or hearing problems, eye sight problems, fatigue and stress, and headaches (Paul & Huws, 2002). Operators at Call Center companies do not have their own workplace; instead, they have to take any available work station that is free. This means that there are higher demands for the furniture and equipment to be adjustable (Bayeh & Smith, 1999). Use of non-ergonomic chairs and desks contribute significantly to the physical health issues (Hannif & Lamm, 2005). The poor ergonomic work station designs are due to the assumption of “one-size-fits-all” by the employers. As reported by Norman (2005), most of the computer equipments are designed for men’s dimensions. This may causes women to work in awkward postures or at higher relative muscle forces. Hannif and Lamm (2005) reported that, fatigue, emotional stress and musculoskeletal problems were the main health hazards among the call center workers. Fatigue was the main issue in outbound call centers than in inbounds and was attributed to the unconventional working hours, long shifts and monotony associated with outbound call centers. Emotional stress was associated with answering the calls from abusive customers, as well as repetitive nature of the work. Thus, studies confirm that, the nature of the work in call centers can lead to a wide
range of health problems. These authors observed that, despite the risks posed by the working environment, the management failed to acknowledge the health risks inherent in the nature of work and call center employees and made little effort to minimize/remedy these problems. The absence of formal hazard reporting and management systems in the call center organizations play a key factor in the attrition rates of the call center workers. Employees and supervisors are unaware of how to go about reporting hazards in the workplace; complaints about such issues may, therefore, not reach management leading them to conclude that these problems do not exist.

2.2 Call centers in India

Call centers are comparatively recent introductions (last 10 years) to the world of career options in India. Following the declaration of liberalization policy by the Government of India (1991), new industries and technologies have mushroomed in the country with multinationals showing greater interest and desire of achieving higher productivity. At present, India occupies the tenth position among the industrialized nations of the world. Call centers are gaining ground in India, especially in the metros and large cities with easy net connectivity to the U.S. and other western countries. It is set to register the highest growth rate in call center services industry in Asia Pacific region. Obtaining reliable figures of call centers in India is problematic, as standard statistics has not as yet listed call centers as a separate entity. The career avenues provided by call centers are one of the best suited for even a fresher. With the opening up of the Indian economy and the advent of globalization, more and more companies from overseas are basing or outsourcing their call center services to India.
Amex and GE pioneered the trend of outsourcing to India by setting up facilities in Gurgaon, near Delhi, in the early 1990s. Some of the other beginners were companies such as British Airways who have also been operating their back office in India since the mid-1990s. By the end of that decade, companies had begun to use call-centers in India for direct voice operations. Low labour costs, skilled and dedicated professional flexible statutory regulations and fairly good connectivity make India the preferred location of outsourcing. The employee strength of call centers in India could vary from 300 agents to 2000 agents/shift. These companies have their own standard guidelines for architecture, interiors and services like air conditioning, fire alarm, electrical, fresh air requirement, etc. It has been reported that in these cases, local architects, interior designers and consultants, find it relatively easy to develop similar facilities in India. However, many Indian builders and developers are leasing out their buildings to new companies or entering into this business to ride the current boom (Iyengar, 2002). With increasing confidence of the companies in the capabilities of Indian operations, higher value added activities such as processing of HR, accounting and other non-core functions are reaching the Indian shores (Batt et al., 2005).

In India, call centers work round the clock handling inbound and outbound calls, web based business, chat, e-mail, and a host of other services. They are broadly classified into international and domestic call centers. Further, they can be ‘in-house’ (serving the customers of the parent company) or sub-contractors (serving the customers of many companies). In India, almost all international call centers function as sub-contractors serving multinational companies while most of the domestic call centers function as in-house service centers. Batt et al., (2005), in their study found that banking is the primary sector served by both international and domestic call centers with the second largest sector being the telecommunications
industry in India. They also found international call centers to be seven times larger (with an average of 741 employees) than domestic call centers (with an average of 104 employees). Proportion of male and female workers varied between the types of call centers. Women workers were found to be higher (60%) in domestic call centers than international (42%) call centers. The use of part-time and temporary workers is extremely low in the Indian context. In India, international call centers employ 99 percent of their workforce as full time hires while domestic call centers employ 92 percent of their workforce as permanent and full time hires.

Indian call center employees need to have a minimum of 15 years education (graduation) with good English knowledge. This is true for both international and domestic call centers. The lowest educational profiles are almost equally distributed among domestic and international centers (an average of 12 years of education- a High School degree), while the highest levels of education are found solely in the international centers (17 years on average- typically a post graduate degree) (Batt et al., 2005). However, even with higher education levels, the call center agents face difficulties in conversing with customers for whom English is the mother tongue. Thus, the depth and complexity with which call center agents can engage with customers is a key issue facing the industry (Taylor & Bain, 2004).

Voice, accent and skills are important attributes of a call center employee in India. Call centers in India are prospering because of the fluent English speaking youth, flexibility and the accent that the Indian community has adopted. By providing little voice-based training, Indians can generate quality services to their employees in United States. As per the internet quotes, big corporate houses tapped this Indian potential of smooth interaction with American and European customers. Call centers in India have effectively developed a niche in
areas of data verification, data capture, tele research, service follow-ups and renewing subscriptions, being the core outbound activities of a call center. Unless there is talented enough manpower to handle such queries, no business organization will progress. Thus, India has taken a step in call center industry ([http://callcenterinindia.blogspot.com/2003_06_01_archive.html](http://callcenterinindia.blogspot.com/2003_06_01_archive.html))

According to Batt et al., (2005), under the heading of voice training, these call centers concentrate on training the ‘tone of voice’, that is, the trainees are instructed to use a polite and positive tone, be clear and calm and allow for pauses. These instructions are given during two hours of training without specific demonstrations or opportunity for practice by the trainees. Call center employees often find it difficult to follow these instructions as different languages and cultures generate different types of customers. The particular style of speaking does not correspond pragmatically to the way service calls are usually conducted in a particular language. The other conflict experienced by the employees is difficulty in keeping ‘positive sound’ intonation on these routine dialogues with reports of intonation patterns flattening out. Cameron (2000) reported that, male CCOs experience difficulty in finding the right tone of voice, as speech style adopted by call centers is symbolically feminine, the ‘smiling’ and ‘expressive intonation’. This leads to the conflict between their construction of masculinity and the requirement of the feminine speech style. This style is difficult for male speakers as they are unused to using the wider pitch range commonly adopted by female speakers. Call center trainees are hence made aware of voice requirements in terms of prosodic features but not in terms of voice use and voice care.
Even though the Indian and western call center environments share many similarities, there are important differences that combine to exacerbate the strain on Indian call center agents. The factors that make Indian agents experience this work as demanding, pressurized and frequently stressful are long travelling time to and from work site, nature of night shifts, not sufficiently developed or flexible English language capability to handle more than a routine call satisfactorily. Along with this, there is neither recognition, nor even embryonic union organization of Indian call centers. This absence of employee voice, denies Indian workers of opportunities to channel their grievances and improve their working conditions. A further source of pressure for Indian call center workers is the requirement that they conceal their location, and even their identities as Indians, by speaking with ‘neutral’ accents and taking on Anglicized pseudonyms, in order to meet the expectations of western customers (Taylor & Bain, 2005). The highly stressful nature of the job has generated high level of turnover among Indian call center agents.

In the Indian context, agents working in call centers are trained to believe that the customer could hear their smiles and sense their moods, and make it mandatory to enthusiastically communicate with the customer, no matter how irate he/she is. They need to learn to use a neutral accent. Call handlers put their emotions aside and attend the next call with equal attention. The evaluation parameters in call centers included the number of calls taken per shift, the average handling time per call and the maintenance of prescribed procedures while in an outbound call center, it was their ability to convert calls into sales. Working for overseas clients located in the US, Canada, UK or Australia meant that they needed to work in night shifts in order to suit the different time zones (Batt et al., 2005).
Excessive work demands in the call centers have led to high turnover as high as 30% (Sudhashree et al., 2005). The BPO employee survey report (2004) suggests that, Indian call centers are ranked high for attrition due to health reasons. There is little evidence on the occupation health and safety (OHS) policies and practices used in call center work places and hence, health and safety of call center employees emerge as one key concern (Taylor & Bain, 1999). Damage to voice caused by intensive use of voice in CCOs has not been fully studied in the Indian set up and calls for a study of this nature.

2.3 Prevalence of voice problems in CCOs

Prevalence is usually discussed in relation to disease and is commonly defined as the number of people with the disease at a specified time divided by the number of people at risk at a specified time (Monson, 1990). Generally, female’s report of higher prevalence of voice problems than males (Fritzell, 1996; Morton & Watson, 1998; Russell et al., 1998; Smith, Kirchner, Taylor, Hoffman, & Lemke, 1998a) attributed to the frequency of vocal fold collision (Vilkman, 2000). That is, frequency of collision plays a role in tissue changes of vocal folds (higher pitched voices are more susceptible to vocal tissue changes than lower pitched voices). Females are thus more susceptible to develop voice problems as a female professional voice user speaking at 250 Hz F0 (5 x 45 minutes) in a working day, would have 1,000,000 vocal fold vibrations, whereas, it will be 50% lower in males. Other reasons quoted in the literature include, difference in concentration of hyaluronic (HA) acid in male and female larynges (Butler, Hammond & Gray, 2001; Chan, Gray & Titze, 2001; Ward, Thibeault & Gray, 2002).
Reliable and valid prevalence data require cross-sectional studies of large and representative population samples (Bowling, 1997). Furthermore, such data would enable analysis of risk factors that may contribute to the prevalence figures. Prevalence of voice problems have usually been investigated within a specific occupational group or among those seeking help from voice clinics. Studies estimating the occurrence of problems throughout all voice users are sparse. The “teaching voice” has been of special interest in several studies conducted in different parts of the world. Besides teachers, extensive voice use is required by call center employees, where the speaker must rely solely on his/her voice without any support from body language or written communication (Lehto et al., 2005). Valid and reliable epidemiological information on voice disorders in call center employees would greatly enhance our ability to plan future health-care provisions. In the Indian set up, information on prevalence (number of existing cases) and nature of dysphonics in call center employees is very limited. It’s felt that accurate prevalence and incidence data are important in elucidating the causes of voice disorders, describing the frequency with which new diseases develop in a population, identifying the characteristics that increase the risk, and developing early screening or disease prevention programs to protect against further deterioration of the larynx and related structures.

According to Titze et al., (1997), in US, the telemarketers are about three times (2.3%) as likely to seek help from a voice clinic as general population, even though they constitute only 0.78% of US workforce. This was remarkably disproportionate to the total work force. Liechavicus (2000), cited in Oliveira, Behlau and Gouveia (2009), assessed the reporting of vocal symptoms in 62 telemarketers using questionnaire and reported the presence of one to three symptoms in 58% of the telemarketers.
Grayson (2001) from Queen Margaret University College, Edinburgh, reported that 78% of workers at one call center had reported of some sort of voice problems of which 60% reported that it affected their work. An association between voice problems and the neck and shoulder problems has been mentioned in the study. Jones et al., (2002) studied the prevalence and risk factors for voice problems in telemarketers in comparison with age and gender matched general population using a self-reported questionnaire. A total of 304 employees completed the survey along with 187 community college students who served as control group. Results indicated that 68% of the telemarketers and 48% of control subjects reporting one or more symptoms of vocal attrition. Multiple logistic regression analysis showed that telemarketers were twice as likely to report one or more vocal attrition symptoms compared to control group.

Taylor & Oates (2004) investigated the prevalence and nature of voice problems in 450 CCOs using self-reported questionnaires. 14% reported the presence of voice problems on the day of the survey and 38% reported the presence of voice problems more frequently (once every couple of months or more frequently during the previous 12 months). When compared to general population, CCOs were almost five times as likely to report of voice problems on the day of survey and almost 15 times as likely to report of voice problems once every couple of months or more frequently. Male CCOs were 20 times as likely to report of voice problems on the day of the survey when compared with their female counterparts in the general population. Equal distribution of voice problems among males and females was reported. This study hence reported of higher prevalence of voice problems among CCOs and called for attention to provide practical strategies for maintaining a healthy voice, and enhance professional performances.
Matic (2006) studied the prevalence of voice loss and voice related problems in Call Centre Agents (CCAs) employed in the Transport for London (TfL) Call Centre and the sickness absence due to voice problems using self-reported questionnaire. The study indicated a significant difference between the reported voice problems in CCAs and the control group with CCAs reporting of poor voice health. CCAs were more likely to have time off for voice related to work when compared to controls. Females in both the groups reported problems with their voice related to work almost twice as often as males.

Vannan et al., (2009) conducted a preliminary study to evaluate the incidence of vocal symptoms among call center employees and prediction of self-voice report for developing a voice disorder using Voice Handicap Index (VHI) and an adapted version of Vocal Self Perception Attitudinal Questionnaire, in India. A total of twenty one call center employees were included in the study (15 males and 6 females) in the age range of 18 - 30 years and were divided into group I (with work experience of 0 - 1.9 years) and group II (with work experience of 2 to 5 years). Their results indicated that 80% of participants in group II and 35% in group I reported of problems in physical aspects and only 20% of group II showed disturbances in emotional aspects. Overall scores revealed that participants of group II reported more voice handicap scores than the group I participants. The attitudinal questionnaire indicated that majority of the participants did not think of their voice. This study is however limited by the small number and cannot be generalized.

Shah and Sanghi (2010) studied the life style, voice problems and reflux symptom index in call center employees in India. They reported that 40% of the international call centre and 30% of the domestic call center workers experienced voice problems and had tea/coffee,
smoke and menthol comparatively more than those working in non-vocal professional set ups. The reflux symptom index values were significantly higher in international call centre employees than the domestic call centre employees and non-vocal professionals. They concluded that call center employees were more susceptible to develop voice problems due to vocal loading and habits such as smoking, gastric reflux and heavy caffeine intake.

The data in the published reports concerning voice problems among CCO’s vary depending on the methods used and how the voice problems and voice disorders were defined. Generally, the results showed that CCO’s were at greater risk of developing voice problems.

2.4 Risk factors for voice problems in professional voice users

Voice problems occur with a change in the voice, often described as hoarseness, roughness, or of a raspy quality. People with voice problems often notice or complain of changes in pitch, voice loss, loss of endurance, and sometimes, a sharp or dull pain associated with voice use. The organs involved in voice production may be affected directly by disease or trauma or indirectly by the condition of parts of the musculoskeletal system which influences respiratory function and by neurological factors which mediate fine muscle control.

The primary risk factors for voice disorders in persons who work in occupations include, vocal misuse behaviours (loud talking, yelling, screaming, hard glottal attacks, singing or speaking outside acceptable physiological range, speaking in a noisy environment, excessive coughing and throat clearing, grunting (as in exercising and lifting), excessive talking, loud, hard, abusive laughing, producing voice when laryngeal tissues are inflamed); exposure to variety of substances (alcohol consumption, medication, caffeine, recreational
drugs, smoke, reflux of stomach contents, inadequate treatment of early symptoms) and psychogenic factors causing increased musculoskeletal tension (Johnson, 1993), work environment factors such as poor room acoustics, poor quality of air, poor working posture, poor quality aids (Sala et al., 2001; Vilkman, 2004; Aronson, 1985; Sataloff, 1991; Stemple, 1995).

Other contributing co-factors include gender, individual endurance, state of health, living habits, vocal skills/experience, and psychosocial aspects/personality. This indicates that causes of voice disorders or factors which make the voice more vulnerable to damage are multi-factorial in aetiology.

A typical call center employee’s profession includes many of the above mentioned risk factors, such as static muscular overload, high number of calls to attend with few breaks, excessive stress, unsatisfactory air quality, and ergonomic conditions (Gilardi, Fubini, d Errico, Falcone, Mamo & Miqliardi et al., 2008). These high vocal demands and unfavourable working conditions may cause problems for the voice in call center employees. The following subsection briefly discusses the above mentioned factors with specific focus on call center operators.

**Vocal loading**

Voice use itself is the basic risk factor for voice problems in professional voice users and can result in vocal loading. The stress inflicted on the speech organs when speaking for a long time is referred to as vocal loading. Professional voice users are more likely to experience vocal loading effects as they have to use their voice to a greater extent than others. The factors contributing to vocal loading are; tendency of a person to habitually speak loud,
excessive, and rapid (Sapir, Attias & Shahar, 1992). Vocational factors such as vocal demands of the job, background noise, room acoustics, speaking distance, air quality, posture, stress, and equipment design are identified as contributing factors for vocal loading (Vilkman, 2000). Voice straining factors such as, speaking at high intensities (more than 70 dB), for long duration, at a high pitch outside the normal range and with strong intonation and abnormal resonance are some of the voice production habits that increase vocal loading (Sodersten et al., 2002). It is generally assumed that persons who have to speak loudly are at great risk for developing throat problems as raising intensity of speech requires raising the sub-glottal pressure with greater glottal adduction. This vocal behavior leads to hyper-functional voice production and vocal fatigue (Rantala & Vilkman, 1999; Baken & Orlikoff, 2000). Yiu (2002) commented that, inappropriate breathing, insufficient rest, certain speaking techniques such as percussive speaking, that is too loud use of voice or stressing the first syllable of each word, could result in muscles getting fatigued. Excessive vocal fold collision (Vilkman, 2004) and amplitude of vocal fold vibrations (Boone & McFarlane, 1983) lead to vocal fold tissue changes. These problems are typically reported by teachers of large class, call center operators working in an eight to nine hour shift, aerobics instructors etc (Vilkman, 2000).

Rantala and Vilkman (1999) found a positive correlation between index of voice loading and subjective voice complaints. That is, the more voice complaints the subjects had, the higher was the value of the index of voice loading. Symptoms such as vocal fatigue, hoarseness, sore throat, throat discomfort or pain were found to be associated with vocal loading (Long et al., 1998; Kooijman, DeJong, Thomas, Huinck, Donders & Graamans et al., 2006). Localized inflammation and swelling of the laryngeal mucosa are commonly seen as vocal loading effects. However, there are individual variations in voice endurance, that is,
some individuals’ tolerate enormous amounts of loading, while others, even despite a good vocal technique, are prone to develop vocal symptoms after moderate loading (Vilkman, 1996). Colton and Casper (1990) stated that, “each individual larynx has physiological limit that varies not only from person to person but also intra-individually as influenced by numerous factors”. Supporting this view, Fawcus (1991), observed that, some people became temporarily hoarse after using the voice against a high intensity background noise, while others did not. This is likely to be based on genetic variations in the biomechanical properties of the vocal folds (Gray & Thibeault, 2002). However, if a speaker constantly uses his/her voice extensively with a poor vocal habit, voice symptoms are likely to occur (Williams & Carding, 2005).

The voice is the most important tool for a call center agent. In telephone communication, the voice is of greater importance because of the exclusion of other communication means (gestures, facial expression). Greater numbers of activities or situations in which call center agents are frequently involved are detrimental for their voice. Although telephone operators are involved in extensive amount of speaking, their phonation consists predominantly of individual communications at low intensities, approximately 55 dB at 1 m, considered to be normal or soft voice (Lehto et al., 2006). Thus, it can be assumed that, in call center operators, the problem is not in terms of loud voice but the duration of vocal use that acts as a risk factor for vocal loading. They get only an occasional five minutes break between talking on the telephone. In some offices, they do not even get to look up the telephone numbers. According to Titze (1999), due to constant vocal fold collision, the epithelial cells may die and shed off. The vocal fold tissue recovery process may take few hours to as much as 72 hours to complete. For CCOs, the accumulated injury during the work
shift can reach a point where day to day recovery is not possible as they are scheduled to work for five days per week with very limited short and long term recovery period. This may lead to some permanent injury which in turn leads to vocal pathology.

Thus, the short term recovery is very poor in call center operators who are continuously engaged in customer service for about 8 hours a day and serve an average 109 customers/day (Batt et al., 2005; Norman, 2005). According to Titze (1999) continuous voice use for two hours is the maximum time allowable, beyond which the vocal mechanism gets strained and tensed. Analysis of the factors of voice loading index also showed that, F0-time is a bigger risk factor for vocal fatigue than variation in F0, indicating that duration of voice use has a significant effect on the development of vocal symptoms (Rantala & Vilkman, 1999). It’s also been reported that the extent of subjective voice complaints correlated positively with the total number of vocal fold vibrations during vocal loading with the general effect being a shift towards hyper-functional or more pressed type of phonation. Vilkman (2004) has reported of vocal loading having significant effects on the fundamental frequency, type and loudness of phonation, vibratory characteristics of the vocal folds and the external frame of the larynx. Loomis (2005) identified extended work time as the main risk factor for the occupational injury. The author reported an association between rate of occupational injury and number of hours worked on a daily or weekly basis. Further, the study indicated that occupational injury among people who worked more than 12 hours per day or more than 60 hours per week had affected 30-40% higher than those working for fewer hours. Working overtime was associated with even higher (60-80%) injury than among people who did not work overtime.
Along with continuous voice use, in order to compensate for the lack of supporting communicative tools such as body language and facial expression, many call center agents speak with a voice and intonation that are unnatural and giving rise to functional disorders (Lehto et al., 2003). This produces enormous vocal load on the call center operator. Cameron (2000) quoted that, “the speech style adopted by call center workers are essentially a gendered, symbolically feminine speech style”. Thus, male employees experience a conflict between their construction of masculinity on one hand and requirement of the feminine speech style. Hultgren (2004) stated that, using positive tone of voice may simply be difficult for male operators as they are unused to using the wider pitch range commonly adopted by female speakers. Call center agents experience a lot of stress from some of their clients which when compounded with vocal overuse, could increase the level of strain and tension within the vocal mechanism and result in exerting more force and effort for vocal production (Alexis, 2006).

From this, it is more appropriate to say that, CCOs are forced into strenuous prolonged voice use which is considered as the main hazard to develop voice problem in professional voice users. These factors are inherent in the working environment and the nature of the profession.

Work environment factors

Posture: The working posture is an essential issue in ergonomics and has also been shown to be connected with voice function as the musculoskeletal system provides support to the respiratory and phonatory system. The entire musculoskeletal system must be considered important for the healthy voice production. Neck, jaw, and shoulder tension are most
prominently recognized factors affecting voice production (Benninger, 1993). Muscular tension related disorders are often noted in professional voice users and lead to vocal fatigue (Koufman & Blalock, 1988).

Literature reports of a relationship between body posture and voice. According to Hoit, (1995) and Kooijman, de Jong, Oudes, Huinck, van Acht and Graamans, (2005), posture affects breathing patterns, especially, which are critical to normal voice production. Angsuwarangsee and Morrison (2002) found a significant relationship between different head positions and increase in laryngeal muscle tension. Increased muscle tension in the neck and larynx can lead to increased muscle tension in the suprahypoid and perilaryngeal muscles during phonation, elevation of the laryngeal level in the neck, an open posterior glottis chink on phonation and associated vocal fold abnormalities such as nodules or chronic laryngitis (Belisle & Morrison, 1983). Vintturi et al., (2001), found significant association between vocal loading and sitting posture. They observed more pressed phonation during vocal loading task in sitting position when compared to standing position. Kooijman et al., (2005) found that, specific combinations of hypertonicity and posture problems are related to high scores on the VHI and low scores on the DSI. This increased muscle tension is maintained due to inappropriate body posture, mental state and personality factors, along with excessive shoulder and neck muscle tension, extended head position, restricted jaw movement. The above mentioned studies confirm the importance of body posture and muscular hyper tonicity in the development of voice disorders.
For the Call Center Employees, work involves in a sitting posture with simultaneous use of a telephone and computer in their work station. As reported by Lehto et al., (2005), call center employees have two harmful working postures. One is the position where the upper part of the back bends down while the shoulders are pushed forwards. In the other, the head leans forward, and causes tension in the neck and throat. These positions require excessive tension in the neck and laryngeal muscles that alter the speaking technique and contribute to voice problems. Glancing position of the head could lead to breathing problems. Consequently, tension is increased in the throat and neck muscles and can cause hoarseness and a variety of symptoms, especially pain and fatigue associated with talking. Thus, muscle tension and diverging postures seem to result in tensed voice production (Kooijman et al., 2005; Thomas, Kooijman, Cremers & de Jong, 2006). Norman et al., (2004) compared the physical and psychosocial working conditions in call center workers and a group of professional computer users from other occupations. They reported that call center operators spend greater time in front of computers with higher proportion experiencing musculoskeletal symptoms than the other group of computer professionals. Hence, musculoskeletal disorders of the upper extremities and neck are the most common occupational health problems associated with CCOs due to inappropriate postures at work and account for majority of work related lost time (Norman, 2005).

**Air quality:** The different factors affecting air quality are humidity, cleanliness, and temperature. Of these, the effect of air humidity plays an important role in professional voice users. According to Kroemer and Grandjean (1997), a relative indoor humidity of 40 - 60% is comfortable whereas, humidity below 30% affects the mucous membranes of the nose and the throat. Dry air and oral breathing increases the phonation threshold pressure and vocal effort
Further, there is an increase in sub glottal pressure in dry environment (humidity level of 30 - 35%) compared to wet (humidity level of 85 – 100%) environment (Verdolini-Marston, Titze & Druker, 1990). Supporting this, a series of studies reported strenuous of voice production associated with vocal symptoms during voice loading tests in laboratory conditions in dry environment (Verdolini et al., 1994; Andrews, 1995; Vilkman, Lauri, Alku, Sala & Sihvo, 1997; Richter, Lohle, Maier, Klieman & Verdolini, 2000; Vintturi et al., 2001). A study on speaking environment (HSE, 2006) reported that, high temperature is usually associated with dry air. Heat generated for 24 hours a day by large number of computers in the call centers increases the risk of low humidity, which can dry air to unacceptable levels, and cause dehydration to sore eyes, voice loss and headaches. Lehto et al., (2005), found that, room temperature was quite high (23° C) and lower percentage of (19%) of the ambient humidity instead of recommended 25 - 45% causing dry air. Jones et al., (2002) reported significant association between telemarketers experiencing dry working environment and their work being affected. These studies indicate that, adequate humidity is essential for better performance of the professional voice users.

Other factors which can affect the air quality are different chemicals or smoke present in the work environment. Chemical agents reported in the literature to be associated with occupational voice disorders are, Freon gas, Formaldehyde, Organic mercury, Diquat Sulphuric acid and Solvents (Tanturri, Pia & Benzi, 1988). Lehto et al., (2003) noticed higher concentration of Formaldehyde (0.1 - 0.2ppm) than the recommended 0.03 ppm at a call center in Finland and opined that this higher concentration of formaldehyde in working environment could be one of the contributing factors for higher rate of sick leave among the call center operators.
Read (2005) reported that, call centers though air conditioned do not have good air quality. Along with this, they are higher density occupancies than other office uses, with air systems not optimized for the employee’s needs. Offices are poorly ventilated with minimal air changes, keeping airborne diseases and toxins inside. Further, high concentration of employees in call centers increases the risk of airborne pollutants and irritants, which in turn, increases the risk of sickness absences, as bacteria and viruses tend to colds and irritate the throat and lungs contributing to the voice problems.

**Room acoustics and background noise:** Noise in the working environment has been considered as relatively well-known risk factor for professional voice users. Many professional voice users need to speak in high background noise levels, which may lead to misuse or overuse of voice and contribute to the development of voice disorders (Vilkman, 2000). Alku, Airas, Bjorkner and Sundberg (2006) reported an association between intensity of speech and the communicative ambience. That is, the speaker increases the intensity of speech in the presence of background noise, poor room acoustics. Different types of background noise at working places are speech babble (such as pupils in a classroom or office workers in an open-plan office), disturbing ventilation or air conditioning noise, or external such as traffic noise. As reported by Webster (1965), noise levels in work spaces should not exceed 62 dB (A). For every 10 dB increase in noise, 6dB increase in SPL of speech is required for talker’s words to be understood for satisfactory communication. Background noise forces the speaker to raise his voice to be heard due to “Lombard effect” (Lane & Tranel, 1971; van Heusden, Plomp & Pols, 1979). Baken (1987) observed an increase in F0, SPL and duration of speech sounds with increased noise. They further stated that, with the aim of improving message transfer in noisy conditions, the speaker increases the vocal fold
adduction which in turn increases the SPL and raises the F0. Along with higher SPL and F0, Sodersten, Ternstrom and Bohman (2005) observed an increase in vocal instability and roughness of the voice, when speaking in higher background noise. They also reported of women requiring to put excessive effort to speak than men in noisy situations. In summary, these studies indicate that, high levels of vocal function due to environmental noise causes a risk to vocal function.

In call centers, the general conversation with callers and colleagues are the main source of background noise. In some, the operators are allowed to listen to their own music, which itself is an additional source of background noise (Patel & Broughton, 2002). Noise from neighbouring or opposite work places and computer cooling fans are also quite disruptive in call centers and restrict the employers’ abilities to concentrate (Benninghoven, Bindzius, Braun, Cramer, Ellegast & Flowerday, et al., 2005). Further, echoes of sounds and noises reflected from hard walls, ceilings, windows and floors contribute to overall noise levels and cause poor room acoustics. The inability to hear due to background noise reduces the speed at which call center employees can attend to customers. With the presence of background noise, call center employees exacerbate the problem by raising the volume on their headsets, and speak louder which further increases the background noise level. Raising the voice can also happen due to poor quality microphones, which in turn increase the noise levels (Jasmine, 2004). However, Patel & Broughton (2002) reported that, the noise levels within the call centers are between 57 and 66 dB (A) even with music playing and well within the permissible limits. Similarly, Lehto et al., (2005) found call center work place background noise level around 42 dB (A). These studies suggest that, the workplace background noise levels are not very high as that of teachers (day care) who usually speak at 72 – 90 dB (A).
Contrary to these studies Benninghoven et al., (2005), reported that, room acoustics, headset acoustics, indoor climate (especially humidity), air quality, lighting and workplace geometry are the important key factors influencing the health of the call center workers. They reported that, improving the room acoustics, optimal humidity (> 45% relative humidity) as well as regular ventilation (CO2 content < 1000 ppm) improves the performance of the call center agents. These authors have reported that, poor room acoustics and indoor climate are the principle factors behind developing voice problems by call center employees.

**Health-related factors**

There are a number of factors which may affect the vocal fold mucosa or the respiratory ability to support voice. The most common factors affecting the vocal mechanism are;

**Smoking:** Smoking is among the most common factor affecting occupational voice users. Smoking leads to vocal fold edema (Reinke’s edema) on the vibrating edge and changes the vibrating pattern, reduces fundamental frequency due to increased mass of the vocal folds, changes the voice quality, nature and capabilities (Marcotullio, Magliulo & Pezone, 2002; Gonzalez & Carpi, 2004). Long and excessive cigarette smoking, chronic industrial or pollutant smoke exposure and inhalation can cause emphysema or shortness of breath. With continued smoking, the mass of the vocal folds increase resulting in changes in the vibratory pattern of the vocal folds (Benninger, 1993). Smoking was found to have a significant association with vocal symptoms in CCOs (Jones et al., 2002; Lehto et al., 2003).

**Hydration:** The viscoelastic properties of the vocal fold mucosa depend on the state of vocal fold hydration (Chan & Tayama, 2002). Vocal fold hydration can be controlled by both
systemic and superficial water content (Fisher, Telser, Phillips & Yeates, 2001; Gray, 2000; Verdolini, Min, Titze, Lemke, Brown & Van Mersbergen et al., 2002). Dehydration leads to elevation in the viscosity and stiffness of the vocal folds. Increased viscosity leads to greater friction and heat dissipation during vocal fold vibration and requires greater energy to initiate and sustain phonation (Verdolini-Marston et al., 1990; Titze, 1994; Sivasankar, Erickson, Schneider & Hawes, 2008; Sivasankar & Leydon, 2010).

Studies have indicated that lowering of the viscosity of the lubricating mucus by adequate hydration is important for adequate vocal fold vibration. This is based on the hypothesis that, hydration treatment may reduce energy dissipation in vocal fold tissues during phonation by decreasing their viscosity. The viscosity of the vocal fold tissue has been shown to be of theoretical importance in vocal fold oscillation. It determines the amount of energy loss due to internal friction in the vocal folds. As viscosity is increased, more energy is required to maintain the same phonatory conditions. Therefore, from the standpoint of vocal efficiency, the viscosity of fluids within the vocal folds should be low enough so that most of the energy provided by the lungs is converted into acoustic energy in the vocal tract rather than heat within the vocal fold tissue.

**Drugs:** Side effect of many medications can be detrimental to the professional voice user. Factors like age, gender, body composition, and metabolism can influence a person’s response to the medication. Nearly every medication has some laryngeal effect, although in most cases, the effects are so minor as to be clinically insignificant (Martin, 1988; Sataloff, Lawrence, Hawkshaw & Rosen, 1993). Many medications (antibiotics, antiviral agents, antihistamines, diuretics, corticosteroids, sprays, inhalants, antihypertensive agents, sleeping pills, hormonal...
medications, bronchoactive medications, beta-blockers, central nervous system stimulants, psychoactive medications) have significant impact on the voice and larynx. Excessive drying of the upper respiratory tract is the vocal complication associated with the largest number of medications by reducing the secretions that normally keep these membranes moist and lubricated. Thick viscous vocal fold secretions lead to aberrations in vocal fold vibration. The drugs used in treating asthma (bronchodilators), upper respiratory tract infections (antihistamines, cough suppressants), hormonal therapy (androgen, testosterone) were found to have effects on vocal mechanism. Decongestant drugs used in cough, cold, sinuses and allergy remedies are known to cause congestion and drying of the mucus membrane (Martin, 1988). Lehto et al., (2003), reported that, the most common risk factor for women working in call centre was the use of hormonal preparation.

In addition, certain recreational drugs, such as alcohol, cigarettes, marijuana, and cocaine have been reported to have significant effects on voice production. All these agents are irritants to the respiratory tract and larynx and could contribute to the development of voice disorders (Stemple et al., 2000). Alcohol is an oral and laryngeal irritant and acts as a vasodilator of the mucosal lining of the larynx. It causes dryness of the mucus membrane of the vocal fold and increases the likelihood of vocal fold haemorrhage during phonation (Sataloff, 1991). Caffeine is also a vasodilator and has dehydrating effect on the mucosal membrane. Excessive caffeine intake can lead to thick, sticky mucus and frequent throat clearing or coughing (Stemple et al., 2000). As per these authors, some of the reasons given by professional voice users for using these recreational drugs include, stress and pressure in the working environment, tight schedules that yield minimal time for relaxation and individual emotional conflicts.
**Allergy/Respiratory illness:** Respiratory illnesses, chronic allergic rhinitis, sinusitis have been found to be associated with vocal complaints (Vilkman, 2000; Long et al., 1998). Allergy is referred to as a hypersensitive or pathologic reaction to environmental factors (Benninger, 1993). Recurrent laryngitis, cough, and post nasal drip could cause chronic inflammation of the membranes that line the nose, throat, vocal folds and trachea. Respiratory allergies can cause decreased pulmonary function, excessive secretions in the lower airway, trachea, bronchi or in the upper airway of the pharynx, edema of vocal folds and unusual resonance characteristics due to congestion of the membrane in those areas (Colton & Casper, 1990; Jackson-Menaldi, Dzul & Holland, 1999). As reported by Boone (1988), allergic reactions or respiratory infections in the upper respiratory tract could produce hoarseness or even a complete loss of voice. These symptoms could cause long term disability and frustration for a professional voice user (Jackson-Menaldi, et al., 1999). Gotaas and Staar (1993), observation that teachers with allergic conditions suffered more from vocal fatigue than unaffected colleagues supports this aspect.

Cohn, Spiegel and Sataloff (1995), found higher incidence of voice disorders in asthmatic than non-asthmatic individuals. Asthma is usually treated with inhalers or oral bronchodilators and or aerosolized steroid inhalers. Prolonged use of aerosol steroids could cause dysphonia (Benninger, 1993; Ihre, Zetterstrom, Ihre & Hammarberg, 2004). Though some studies believed in the presence of an association between sinusitis and dysphonia, the pilot study conducted on 19 chronic sinusitis subjects by Cecil, Tindall and Haydon (2001), did not find sinusitis as a risk factor for dysphonia. Lehto et al., (2003) observed that, frequent colds, allergy or asthma were most common risk factors for the development of vocal symptoms in CCOs.
Gastroesophageal reflux (GERD): The backflow of the gastric contents into the oesophagus is termed as Gastro Esophageal Reflux (GER) and the gastric content travelling up to the level of larygo-pharynx is termed as Laryngo-Pharyngeal Reflux (LPR). In recent years, it has been recognized that, reflux disease or back flow of stomach contents is not just confined to the oesophagus but can reach the larynx and hypo-pharynx have effects on other regions of the upper airo-digestive tract such as the trachea, sinuses and middle ear. These manifestations are believed to be caused by direct contact of the gastric content and injury to the pharyngeal or laryngeal mucosal surfaces (Koufman, Sataloff & Toohill, 1996; Belafsky, Postma, Amin & Koufman, 2002; Sataloff, 2008; Lowden, McGlashan, Steel, Strugala & Dettmar, 2009). As reported by Powitzky (2002), LPR leads to laryngeal pathology such as vocal nodules, Reinke’s edema, scar formation as in idiopathic subglottic stenosis, functional laryngeal movement disorders such as muscular tension dysphonia, paradoxical vocal fold motion and paroxysmal laryngospasm.

Lehto et al., (2003) identified gastro-oesophageal reflux as the common risk factor associated with vocal symptoms in CCOs. Jones et al., (2002) found that, telemarketers who experienced acid reflux indicated their work being affected and which was statistically significant.

General health: Professional voice users are considered as vocal athletes, and their strength, endurance, and fine motor skills necessary to optimize voice production are similar to any athletic endeavours, to work effectively (Benninger, 1993; Vilkman, 2000). In addition to the direct effect these conditions have on voice, it is likely that the speaker develops further compensatory behaviours that may worsen the voice problem (Williams & Carding, 2005).
Eating disorders such as anorexia nervosa and bulimia may be associated with voice dysfunction (Benninger, 1993). Excessive stress and intake of certain medications, such as steroids, could secrete the gastric acid even without intake of food. Similarly, large meals with high fat contents delay gastric emptying. Presence of these can increase the chance of laryngo-pharyngeal reflux (Hanson, Jiang & Chi, 1998). Sleep is an active physiologic process. Regimented sleep-wake cycles may be disturbed in professional voice users, and sleep deprivation can result in excessive physical and emotional strain which interferes with their performance (Benninger, 1993). Rosen and Sataloff (1997), pointed out that, a good night’s sleep is an essential factor for appropriate vocal production, and sleep deprivation can affect motor reflex and co-ordination, reduce resistance to stress, and generate depression and anxiety. Lack of adequate sleep may impair perception and cognition, decrease concentration levels and increase difficulty in performing work related tasks (Krueger, 1989). Ferreira et al., (2010), found that vocal disorders such as vocal fatigue in teachers were significantly associated with less than six hours of sleep, and waking up feeling replenished. Lack of enough rest and sleep leads to whole body fatigue, and the vocal mechanism in CCOs. The functioning of the larynx is, thus, compromised in this condition and makes it prone to overuse (Alexis, 2006). Proper diet and body weight are very important in maintaining fitness of the body. Muscle glycogens are the chief source of energy, and important during moderate to high levels of activity for prolonged periods of time. Glycogen synthesis is directly related to carbohydrate intake (Benninger, 1993). Ferreira et al., (2010), found significant association between lack of regular meal schedule and symptoms of vocal fatigue in teachers.

According to Noronha and D’Cruz (2006), call center employees, on shifts are provided with two 15-minute breaks and a break of half an hour for lunch/dinner, the timing
of the breaks are decided by the team leaders on the basis of call levels or call queue. This non-certainty of time are forced them to choose fast food or skip meal/snack in order to log back in time. This resulted in digestive problems among call center employees. Sudhashree et al., (2005), reported Burn Out Stress Syndrome (BOSS) as the most commonly seen condition among Indian call center agents characterized by chronic fatigue, insomnia and complete alteration of biological rhythm of the body. The frequent consumption of coffee probably to avoid feeling sleepy and to be awake is another known habit of call center employees. Coffee, however, causes a person's throat to dry and can also give a feeling of something lodged in the throat. This feeling necessitates a person's frequent throat-clearing, a vocal abusive behaviour (Alexis, 2006).

According to the data quest BPO employee survey (2004), call centre employee attrition rate is very high due to health problems in India. The survey reports that, 83% of CCOs are suffering from sleeping disorders compared to the industry average of 39.5%, 8.5% of the call centre employees experience voice problems against 3.9% of the industry average. In general CCOs perceived their voice problems as being associated with their general health status (Taylor & Oates, 2004). Noronha and D’Cruz (2006) reported that most of the call center operators spend their salary on paying off doctors as their immunity system has suffered.

**Hearing loss:** Auditory feedback is critical to allow the individual to monitor his/her sound production, refine pitch and volume, and blend with surrounding accompaniment or voices. Noise induced hearing loss is the most common hearing disorder among the professional voice users. Chronic exposure to loud noise can result in temporary threshold shift or
permanent threshold shift (Benninger, 1993). According to Lane and Tranel (1971), speakers who have hearing difficulty have less control over their voices, and use higher pitch and increased vocal fold adduction, resulting in vocal loading. Study by Gotaas and Staar (1993), reported that teachers with hearing loss suffered from vocal fatigue. According to the reports of Health and Safety Executive (2006), CCOs get exposed to prolonged exposure to noise through their headsets causing them tinnitus and hearing loss. Some reported these symptoms for shorter period while others, of permanent discomfort or impairment. However, Patel & Broughton (2002), indicated that, the risk of hearing damage from using headsets was extremely low and that CCOs are occasionally exposed to high noise levels for short periods with levels unlikely to exceed 80 dB(A).

**Psychosocial and psycho-emotional factors**

Psychological factors play a part in many aspects of human behaviour as they relate to vocal performance. There is a direct relationship between stress and vocal quality. Under extreme stressful conditions, vocal folds may be free of visible pathology, but their precision of movement may be adversely affected by the state of mind and the emotions of the owner of the voice (Stemple et al., 2000). Along with this, some inherent personality characteristics and anxiety could aggravate the voice problem (Boone, 1991; Mattiske et al., 1998; Wellens & van Opstal, 2001). The increased psychological factor increases musculoskeletal tension, pain and discomfort in the laryngeal muscles causing hoarseness, reduction of pitch range, poor vocal focus due to tight laryngeal/pharyngeal constriction that contributes to vocal problems (Johnson, 1993). Boone (1991), identified, breathy and double voice, dry mouth and throat, harshness, high pitch, hoarseness, lifting up larynx, loud voice, low pitch, monotone, aphonia, neck or throat pain, pitch breaks, shortness of breath, strained voice, throat clearing, tight
voice, traumatic laryngitis, voice breaks and weak voice as some of the vocal symptoms associated with stress.

Stress is more inherited in certain occupations than others (Boone, 1991). Titze (1994) stated that professionals such as sales people, teachers, actors, singers, lawyers, politicians, telephone operators, ministers, speak for long hours and were often under considerable psychological stress. Typically, these professionals abuse their vocal mechanism under considerable psycho-emotional stress experiencing tensions in head and neck muscles. Stress is said to reduce behavioural efficiency and disturb voicing. Voice production under stressful conditions could lead to breathy and strained voice quality. Stress has been associated with lower fundamental frequency, intensity and aerodynamic capacity (Lierde, Heule, De Ley, Mertens & Claeys, 2009).

There is now an almost universal consensus that call center work is stressful, and obviously a significant portion of the call center literature is devoted to detailing the sources of stress. The primary source of stress has been reported to the job nature inclusive odd timings and job nature involving seating for 9 hours a day, reading prepared text over the phone for a long duration and communicating often with irritated and impatient customers. (Holman, 2003; Sudhashree et al., 2005; Rocha, Glina & Nakasato, 2005). Low levels of control in the job, fixed performance goals (number and length of calls), the presence of performance monitoring systems, shortage of breaks and difficult client interactions have been described as some of the other main causes of stress among these workers (DiTecco, Cwitco, Arsenault & Andre, 1992; Deery, Iverson & Walsh, 2002; Holman, 2003; Suri & Rizvi, 2008). Specifically, in telephone interaction, there is a demand for co-ordination of
optimum psychological, behavioural and environmental settings to maintain an efficient balance for the interaction (Cameron, 2000). Thus, a subtle paralinguistic vocal change may unconsciously affect the call-agent, and also the receiver negatively.

In a study by Jones et al., (2002), those telemarketers who frequently experienced stressful calls were significantly more likely to have indicated that their work including voice use was affected. Lehto et al., (2006) suggested that, when the call-agent's voice becomes strained, hoarse or effortful, this in turn adds up physiological and cognitive stress which can result in less efficient interaction, particularly when following a script or protocol for calls. Lin, Chen, Hong and Lin (2010), examined the patterns of major job stressors in call center workers and found 54% of the participants reporting ‘dealing with difficult customers’ as the most significant job stressor.

Understanding the effect of stress on voice production is relevant for clinical diagnostic situations, particularly in patients who present with a functional (non-organic) voice disorder, as well as those who have some organic pathology and also report exacerbation of symptoms during periods of increased life stress. It is hence clear that, call center workers’ life style and work expose them to many health related problems. These health related factors can lead to voice problems. The main risk factors which could contribute to the voice problems in call center operators are excessive use of voice, limited vocal recovery time and work related stress. The effects of over use of the voice are further compounded by background noise, poor room acoustics, inappropriate working postures, poor atmospheric humidity.
2.5 Vocal symptoms in professional voice users

Generally, vocal symptoms start slowly and increase gradually over time in professional voice users. They develop from sporadic to permanent with the appearance of organic lesions, especially vocal nodules (Tavares & Martins, 2007). Behavioural, biological and or psychosomatic factors significantly influence the development of vocal attrition symptoms in professional voice users (Sapir et al., 1990). Each individual larynx has physiological limit that varies from person to person (Colton & Casper, 1990). That is, the presence of vocal pathology may induce vocal complaints only in some and not in all (Heman-Ackah, Dean & Sataloff, 2002). Casper and Murray (2000) indicated unintentional change in the pitch, vocal fatigue, and frequent sore throats as the early warning signs of vocal attrition. Vocal fatigue is the most common hazard experienced by the occupational voice users (Koufman & Blalock, 1988). Vocal fatigue is described as a negative sensory vocal symptom that corresponds to a change in vocal quality, deterioration of vocal control, discomfort in voice production areas and/or vocal limitations (Boone & McFarlane, 1988). The voice qualities experienced by those suffering with laryngeal fatigue vary with dysphonia progressively worsening towards the end of the speaking day. The fatigued voice may sound hoarse in the ears of the listener or to the speaker him/herself. Symptoms of vocal fatigue include; voice production felt to be more laborious, unusual effort to talk, fullness in the throat, pharyngeal and laryngeal dryness, weak voice or inability to produce voiced sound. The speaker may also have symptoms such as an urge to clear the throat, a feeling of mucus or a lump in the throat (Sapir, 1993; Gotaas & Starr, 1993; Rantala & Vilkman 1999; Vintturi, 2001).
Vocal fatigue can be central fatigue (tiredness without vocal symptoms) or peripheral fatigue (changes in functions of the larynx and vocal mucosa). Hence, the quality of voice varies on a continuum between normal and abnormal extremes. Thus, the differentiation between a pathological voice and a normal one is complicated (Treole & Truedeau, 1997). Speakers who have no pathology in the larynx may produce hoarse voice and people with certain laryngeal pathologies, on the other hand need not necessarily reveal appreciable vocal disturbances until a considerably advanced stage.

Among the telemarketers, according to Jones et al., (2002) most frequently reported vocal symptoms were; dry and scratchy throat, need to clear throat, tired or weak voice, voice breaks, increased vocal effort, loss of voice at the end of the sentences.

Lehto, et al., (2003) studied the voice symptoms, associated risk factors, and effect of vocal training in 38 female and 10 male call center operators. The most common symptoms (once a month or quite often) reported by females were feeling of mucus in the throat, and need to clear the throat. The problems that were experienced ‘a few times a year or sometimes’ were hoarseness without cold, vocal fatigue after extensive speaking, sensation of dry throat, worsening of voice quality during the day and feeling of lump in the throat. Throat clearing was the most commonly reported symptom among males (once a month or quite often). Other symptoms reported were voice fatigue and sensation of dry throat. The severity of hoarseness was more in females compared to males.

Taylor & Oates (2004) reported that the most frequently reported vocal symptoms by the CCOs were vocal fatigue, poor voice quality, voice breaks and increased vocal effort. The most commonly reported throat discomfort symptoms were: dry throat, scratchiness or
tickling sensation and frequent need to clear throat or cough. They also identified talking loudly, coughing or clearing throat, talking excessively and talking with poor posture as the most frequently reported vocal misuse behaviours by the CCOs.

Sudhashree et al., (2005) reported that around 8.3% of call center operators complained of voice loss. The symptoms were characterized by inability to speak, pain in the throat, croakiness of voice, irritation and cough, poor vocal power and breathing difficulties.

Lehto et al., (2005) studied the vocal symptoms of telephone workers at four different moments of the working day (morning, before the lunch break, after the lunch break and end of working day) and the effect of short vocal training on 35 female and 10 male customer service advisors. Their results indicated a linear increase in all subjective voice symptoms in customer service advisors during working day. The most prominent symptoms were hoarseness, feeling of vocal strain and a feeling of mucus in the throat. Of these, the feeling of vocal strain invariably increased throughout the day. Increase in hoarseness was reported both during the lunch break and in the afternoon. Increase in symptoms in customer care advisors was more after lunch compared to before lunch. Increase in the amount of mucus was during the lunch break. Before lunch, two out of eleven symptoms were reported to be increased, while nine out of eleven symptoms increased after the lunch break.

Lehto et al., (2006); (2008), evaluated the experience of subjective vocal symptoms before and after working day in customer service advisors. The subjective vocal symptoms such as vocal fatigue and hoarseness increased significantly for females and hoarseness for male customer service advisors, at the end of the day; this indicated the time of the day as having significant effect on subjective symptoms. The phoniatric examination of these
participants showed normal findings in most of them except few having glottal edema and irritation and who were smokers.

Lin et al., (2010), assessed the health status of the inbound and outbound call center employees and found that musculoskeletal discomfort, hoarse or sore throat and eye strain were the most prevalent health complaints in the telecommunication services. They further reported that, inbound service employees displayed higher prevalence of these symptoms than outbound services. They also found that, outbound service employees experiencing less job demands and high work related support than inbound service providers.

On scrutiny of the different studies conducted on CCOs, it can be concluded that, vocal strain, vocal fatigue, hoarseness, poor voice quality, voice breaks and increased vocal effort, dry throat, scratchiness or tickling sensation and frequent need to clear the throat are the most frequently reported vocal symptoms in CCOs.

2.6 Impact of voice disorders in professional voice users

The professional voice users are directly dependent on their voice for livelihood. The impact of voice problems on this population is twofold. Voice problems in this population not only cause vocal symptoms that are characteristic of the disorder, but also carry with it, a high level of emotional strain and anxiety. This anxiety is caused by the disorder’s potential impact on the person’s reputation, the ability to meet professional commitments, or simply the ability to perform his or her job. These concerns and anxieties add to the actual causes of the voice disorder (Stemple et al., 2000). According to several studies published in the late 20th century, the proportion of professional voice users who suffer from voice problems at least a few times a year varies from 20 to 80% (e.g., Russell et al., 1998; Pekkarinen et al., 1992). Voice
problems in practicing professional voice users are particularly imperative as their primary job functions were severely affected (Titze et al., 1997). Moreover, their voice disorders would also affect their social, psychological, and communication functioning adversely (Ma & Yiu., 2001). The measurement of functional impact relates to the degree of impact a disorder has on an individual. The impact of voice problems on an individual does not depend merely on the severity of the symptom but also on how an individual perceives, reacts, and adjusts to it (Yiu, 2002). Thus, understanding the significance of a seemingly straightforward symptom, hoarseness or dysphonia, must include the patient’s needs for voice use and how the dysphonia will alter his/her physical, social, and emotional well-being (Kitch & Oates, 1994; Murry, Medrado, Hogikyan & Aviv, 2004). According to Behrman et al., (2004), perception of vocal handicap or activity limitation due to voice problems depends on the ability of the individual to adapt, compensate for and overcome voice disorders. These features are found to be highly individual and to a certain degree, free of voice use. When the VHI scores in 100 patients with benign mucosal vocal fold lesions were evaluated, professional voice users had higher VHI scores as compared to routine voice users. However, it was also noticed that, singers and actors are more sensitive to smaller voice changes and seek early medical attention.

Smith et al., (1997), studied the impact of voice problems on quality of life in treatment seeking adults. Their findings indicated markedly greater functional problems due to voice in the treatment-seeking group compared with control subjects. In these patients, voice problems negatively affected their social interactions leading to social isolation, depression, and negative professional self-esteem. Their study also observed that, voice problems had negative impact on intelligibility of speech, communication difficulties in
background noise, telephone conversation. In the literature, impact of voice problem has also been measured in terms of number of days missing work (Russell et al., 1998).

Jacobson, Johnson and Grywalski (1997) reported that, degree of self-perceived voice problem is related to the degree of perceived functional deficits. They used the Voice Handicap Index (VHI), to address the impact of voice problems on dysphonic individuals and found that, patients who perceived their voice problems as ‘‘mild’’ indicated relatively mild functional deficits in intelligibility, social activity, job function, emotional status due to voice, and mild to moderate physical problems associated with voice production. Patients who perceived their voice problems as ‘‘severe’’ indicated considerably worse problems in the same domain. This finds support in the study by Murry and Rosen (2000), who also reported that, the degree of to which a voice disorder impacts an individual’s day to day activities may vary significantly depending on the severity of the voice disorder and the voice needs of the patient.

Krischke, Weigelt, Hoppe, Kollner, Klotz, Eysholdt and Rosanowski (2005), studied the health related quality of life (HRQL) in male and female subjects having organic and functional voice disorders. They found that voice disorders had significant influence on patients’ HRQL. They however, did not find any significant correlation between HRQL and kind of voice disorder (functional or organic) or gender. Bassi, Assuncao, de Medeiros, de Menez, Teixeira and Gama (2011), assessed the impact of voice problems on the quality of life in teachers and reported no significant relationship between otolaryngologists’ diagnosis or grade of dysphonia with quality of life.
Impact of voice problems in telemarketers/CCOs has not been studied extensively. Jones et al., (2002) found, 31% of the telemarketers reporting of voice problems affecting their work productivity by an average of five vocal attrition symptoms and half of them missed work owing to their voice problems. Taylor and Oates (2004) reported that, CCOs had reduced social interaction and taking time off work due to their voice problems. Lehto et al., (2003) reported that voice problems in CCOs caused increased sick leaves.

Thus, from the above mentioned studies, it is clear that, voice problems have obvious economic consequences and create negative effects on their occupational performance and society. A shared understanding of voice and shared terminology are important between voice experts and voice users, as the latter judge the success of the treatment or training according to their own perception of their voice improvement, regardless of how the experts measure the treatment or training response (Sellars & Dunnet, 2002; Lee, Drinnan & Carding, 2005)

2.7 Objective and Subjective measures of voice in professional voice users

Self-reported voice problems, acoustic measures, physiologic measures and perceptual evaluation by trained listeners offer different perspectives on describing vocal function. Self-report of voice problems is an useful method for establishing the extent to which the individual suffers from vocal dysfunction. Patients’ self-reporting of voice problems are subjective and depend on their moods and other intrinsic and extrinsic factors as well as their illness perception. Voice is considered to be normal until it can serve the individual’s daily needs (Sapienza & Woodson, 2009). The structural, physiological, acoustic and auditory perceptual analyses are voice evaluation procedures used to quantify the severity of voice disorders. Kent (1994) stated that “a comprehensive assessment of speech function depends
upon a balance of physical and perceptual analyses. Exclusive reliance on either one alone may limit the understanding of speech impairments”. It is unlikely that a strong correlation will exist between quantitative and qualitative measures as instrumental measures are focused on one component of the voice signal (frequency or intensity) while perceptual evaluation is influenced by multiple factors. Thus, it is prudent to consider them separately.

**Acoustic voice analysis findings in professional voice users**

The acoustic signal is the byproduct of phonation (the oscillation of the vocal folds as determined by aerodynamic and myoelastic forces). Perkins (1971) stated that “the underlying premise for detection of laryngeal pathology is that a deviant condition will result in an acoustic “signature” affecting fundamental frequency, intensity, or quality, singly or in combination. Acoustic analysis software’s are able to trace sound waveforms by processing signals and applying algorithms. Analyses of the fundamental frequency, perturbation measures such as jitter and shimmer, and measures of noise make it possible to describe the human voice almost completely (Finger, Cielo & Schwarz, 2009). Haynes and Pindzola (1998) indicated that acoustic measures may be useful in the early detection of vocal pathological conditions, even though no visually detectable lesion or tissue changes are present. Hence, there is a great deal of correspondence between physiology and acoustics, and much can be inferred about the physiology based on the acoustic analysis (Colton & Casper, 1996). On the contrary Behrman and Orlikoff (1997) reported that, the relationship between the physiology and acoustics is not perfect as the voice signal is a complex product of the nonlinear interaction between the aerodynamic and biomechanical properties of the voice production system. Hence, accurate predictions regarding underlying phonatory physiology cannot be made on the acoustic signal alone.
For the acoustic analysis, in the area of occupational voice, speech researchers traditionally use sustained vowels. However, there has been a debate on whether analysis of a sustained vowel is a sufficient representation of human voice communication. It is often presumed that a monophthong vowel, maintained at a comfortable pitch and loudness, should represent the speaker's most stable vocal performance (Heiberger & Horii, 1982; Linville & Korabic 1987). According to Karnell (1993) and Scherer, Vail and Guo (1995), sustained vowel samples are the most reliable speech tasks for acoustic analysis and should have sufficient segment length to preserve the validity of acoustic measurement. The degree to which the acoustic characteristics of a steady-state phonation can be held constant is thought to provide important insight into vocal function.

Some researchers claim that sustained vowels lack many of the properties that running speech has (Klingholz, 1990; Qi, Hillman, & Milstein, 1999). However, the analysis of continuous speech sets higher requirement on the speech analysis than the sustained vowels (Baken, 1987). The most obvious difference between sustained vowels and continuous speech is that the latter is a combination of different utterances and phonemes, and they alternate in rapid succession. An acoustic assessment of phonation involves extracting precise information from a signal that is highly influenced by vocal tract status and function. This task is particularly difficult during speech production when supraglottal behaviours, as well as ventilatory and laryngeal activity, are continuously adjusted. Thus, acoustic material for the evaluation of voice quality is almost exclusively drawn from maximally stable vowel prolongations to avoid the effects of speech intonation and interactions between the larynx and the vocal tract (Baken, 1987). For the acoustic assessment of phonation universally, vowels /a/, /i/, and /u/ are more widely used, as vast majority of the world’s languages contain
these three vowels. This near phonetic universality makes data cross-comparison using these
vowels more feasible across clinics and countries giving access to the intrinsic quality of a
voice without the influence of either consonantal structures or language (Maddieson, 1984).

There are several basic acoustic measures that are part of a standard acoustic analysis.
Multi-Dimensional Voice Program (MDVP) was developed utilizing the Computerized
Speech Lab (CSL) (Kay Elemetrics Corp.). CSL, a hardware/software system which uses an
MS-DOS based computer as host, includes signal conditioning, 16-bit A/D converters, dual
digital signal processors (DSP16A & TMS32025) and support peripherals. The MDVP
system computes a set of 33 acoustic voice parameters in about 16 seconds and provides
flexible routines for graphical representation of the results. MDVP acoustic analysis enables
detailed analysis of 33 vocal parameters viz., of fundamental frequency, perturbation
measures (frequency and amplitude), spectral energy balance and presence of any breakdown
in sonority and diplophony (Nicastri, Chiarella, Gallo, Catalano, & Cassandro, 2004). MDVP
acoustical analysis of voice is regarded as an objective and reliable measure of voice.

Different types of vocal pathologies may have direct influence on these acoustic
parameters. The presence of vocal pathology can increase the turbulent noise in the voice
signal, frequency and amplitude irregularities, variations in sub-harmonic frequency
components, voice breaks and tremors. Understanding the acoustics of these changes is the
key factor in the evaluation of pathologic voices (Kent, 1993). According to the reports of
Baken (1996), acoustic analysis has the potential benefit of measuring and quantifying subtle
differences in voice quality more reliably than most perceptual measures. Normal and
pathological voices were shown to differ in various acoustic measures, such as fundamental frequency, amplitude- and frequency- perturbation and different signal-to-noise indices.

**Fundamental frequency**

The frequency of glottal vibrations per second determines the fundamental frequency of speech, which is commonly denoted by F0 and expressed in Hertz (Hz) or cycles per second. The pitch is the perceptual counterpart of F0. The fundamental frequency is the most saliently used parameter in voice analysis research. The mean fundamental frequency can be measured during sustained vowels or extracted from connected speech. Changes in habitual frequency or frequency range are considered to be important indicators of increased vocal flexibility, loss of mass on the vocal folds, or more appropriate pitch placement (Stemple et al., 2000). Many studies have reported an increase in F0 due to vocal loading (Stemple 1995; Vilkman, Lauri, Alku, Sala & Sihvo, 1999; Rantala et al., 2002; Sodersten et al., 2002; Schneider, Enne, Cecon, Diendorfer-Radner, Wittels & Bigenzahn et al., 2006). Rantala and Vilkman (1999) reported of higher F0 after vocal loading to imply a higher mechanical load posed on the vocal folds. A large increase in F0 followed by vocal loading can be correlated with tiredness of throat. Raising F0 may be used to compensate for loading changes of the vocal fold tissue. It’s also reported that, if the habitual F0 is already higher before loading effects, further increase in F0 during loading task can cause more symptoms of vocal fatigue.

Even though it is a common assumption that, rise in F0 associated with vocal loading, not all studies however, have found significant increase in F0 as a result of vocal loading (Neils & Yairi, 1987; De Bodt, Wuyts, van de Heying, Lambrechts & Abeele, 1998; Jonsdottir, Laukkanen & Vilkman, 2002). Apart from this Artkoski, Tommila and Laukkanen
(2002) found increase in F0 during a day without special vocal load indicating increase in F0 as a normal and healthy adaptation to a situation.

**Perturbation measures**

Perturbation is defined as the cycle-to-cycle variability in a signal. Two common expressions of this term are jitter (frequency perturbation) and shimmer (amplitude perturbation). Perturbation values describe the stability of voice, jitter reflecting frequency, and shimmer, amplitude changes in sequential periods. Perturbation measures reflect the speaker’s ability to maintain respiratory and laryngeal capacities to produce a stable voice output (Baken & Orlikoff, 2000). Every act of phonation is characterized by a certain degree of apparently random variability of the cycle-to-cycle duration (jitter) and or amplitudes of the pitch periods (shimmer) of the laryngeal waveform. Acoustically, perturbation can characterize the laryngeal waveform in both the time domain and the frequency domain. In an auditory perspective, jitter and shimmer contribute to the roughness. Severe perturbation is almost always sign of either pathological or functional disorder, but slight perturbation is evident in all speaking voices and could be attributed to abrupt transitions occurring during voicing onset and offset. These changes in the vocal folds lead to momentary irregularities of phonation before the system settles to a more consistent pattern (Laver, Hiller & Beck, 1992). A perturbation difference of the cycle-to-cycle period in the normal voice is very small and less than 0.5 ms and is not audible (Liberman, 1963). There is a range of mathematical calculations for these terms, which are usually specified by three general parameters; length of the voice analysis window (short-term or long-term averages), absolute or relative measurement units, and statistical emphasis (means and coefficients of variability). Perturbation measures can be difficult to interpret because of the variable potential...
contributors to vocal instability such as aerodynamic, neurologic, biomechanical, source-filter factors that affect the cycle to cycle variability of acoustic frequency and amplitude (Titze, Horii & Scherer, 1987; Higgins, Netsell & Schulte, 1994).

**Jitter:** Measures of cycle-to-cycle variation in fundamental frequency, or jitter, are of interest as an objective means of differentiating normal from disordered voices and of quantifying the severity of vocal deviation (Karnell, Scherer & Fischer, 1991). The normal voice contains some amount of irregular components due to chaotic nature of the laryngeal mechanism, and voice without irregularity is not perceived as human (Titze & Winholtz, 1993). For normal speakers, contributors to jitter are thought to consist of a wide range of factors including variability in vocal fold mass, distribution of mucus on the folds, symmetry of vocal fold structure, timing of laryngeal muscle action potentials, and glottal flow characteristics (Titze et al., 1987; Higgins & Saxman, 1989). Jitter tends to increase during vocal pathology and is responsible for the perception of a harsh, hoarse, or rough voice quality. A high degree of jitter results in a voice with roughness that is usually perceived in recordings of pathological voices. Therefore, a reliable estimation of jitter can be used to discriminate between healthy and dysphonic speakers (Kreiman & Gerratt, 2005). As a general rule of thumb, a mean cycle-to-cycle period difference under 100 microseconds (0.0001 s) is characteristic of a controlled, sustained phonation produced by the normal speaker. However, research has indicated that this absolute jitter is influenced by the mean F0 of the speaker’s phonation (Orlikoff & Baken, 1990). For this reason, several relative jitter measures have been proposed. It is generally expected that normal voices are associated with an F0 variation less than 1% of the mean phonatory frequency (or period). Thus, human phonatory system is not a very stable mechanism and every speaker’s vibratory cycles are erratic to some extent. So, an abnormal
larynx should produce more erratic voice than a healthy one. Even though, literature reports of increased vocal jitter were clearly associated with voice disorders, some have questioned the sensitivity and specificity of these measures (Zyski, Bull, McDonald & Johns, 1984; Ludlow, Bassich, Connor, Coulter & Lee, 1987). Supporting this view, Ng, Gilbert and Lerman (1997), reported that, jitter was not a sensitive acoustic parameter to laryngeal changes associated with acute laryngitis, while, Bonilha and Deliyski (2008), found acoustic jitter and shimmer levels above the normative threshold of MDVP indicating increased perturbation measures even in the normal voice samples. Thus, at present, insufficient normative data impede the use of jitter for clinical purposes.

**Shimmer:** Shimmer or intensity perturbation is a measure of the fluctuation in the amplitude of a sound signal from vibratory cycle to vibratory cycle. It can be expressed in percentage or in decibels (Stemple et al., 2000). It is measured best during sustained phonation of a vowel. It signifies the stability of the signal. Shimmer values serve to quantify short term amplitude instability that does not alter the qualitative features of the vocal waveform (Titze, 1995). The contribution of the amplitude perturbation to identify the specific abnormalities of the glottal function is not clear (Baken & Orlikoff, 2000). Still, shimmer measures remain popular among the clinicians, and are included in most automated voice analysis systems.

Some authors (Peterson, 1958; Baken & Orlikoff, 1988) consider the amplitude variations as the most significant factors in determining the severity of the phonation disorder. Intensity reflects the changes in subglottic pressure brought about by co-ordination of both laryngeal and respiratory muscles and elastic forces. In pathological cases, the vibratory cycle can take place without complete closure of the glottis. This lack of complete closure creates a
breath which compromises the ability to produce a constant sound emission and leads to changes which can affect the vibratory amplitude without affecting the frequency. The vibratory amplitude can also be affected in vocal fold hyper-adduction, which is reflected in short-term control (Nicastri et al., 2004).

Mendoza and Carballo (1998) analysed the effects of stress on acoustic parameters of voice. They observed significant increase in fundamental frequency, jitter, shimmer, high frequency harmonic energy, and increase in spectral noise during experimentally induced stress. They also commented that, the response to a stressful stimulus demands a high level of the vocal muscles, producing a tense voice with degree of tension varying on the perceived stress. Increase in tension can interfere with the vibratory regularity of the vocal folds, resulting in a more aperiodic vibration with possible increase of jitter and shimmer.

**Noise related parameters**

The different acoustic measures which can give information related to the spectral noise are; Noise to Harmonic Ratio (NHR), Voice Turbulence Index (VTI), and Soft Phonation Index (SPI). Acoustic analysis programs often produce a ratio measure of the periodic or harmonic signal energy to the aperiodic or noise energy in the voice waveform. The general principle is that, greater signal or harmonic energy in the voice reflects better voice quality and large noise energy (random aperiodicity in the voice signal) represents more abnormal vocal function (Titze, 1995). Presence of noise components in the main formants of various vowels, high frequency noise components and loss of high frequency harmonic components are described as the acoustic characteristics of hoarse voice and more noise
replaces the harmonic structure as the degree of perceived hoarseness increases (Yanagihara, 1967).

Harmonics-to-noise ratio also known as noise-to-harmonics ratio and signal-to-noise ratio is a general evaluation of the presence of noise in the analyzed signal (including amplitude and frequency variations, turbulence noise, sub-harmonic components and/or voice breaks). NHR is a ratio of the in-harmonic energy in the range 1500-4500Hz to the harmonic spectral energy (70-4500 Hz). This is usually expressed in decibels. Lower HNR indicates noisy or hoarse voice and high HNR indicates a relatively clear voice. The degree of hoarseness can be evaluated by judging the extent to which noise replaces the harmonic structure in the spectrogram. ‘H’ is the energy of the averaged waveform, while ‘N’ is the mean energy of the differences between the individual periods and averaged waveform. The ratio of H/N is expressed in decibel scale and a ratio of lesser than 7.4 dB is considered pathological (Yumoto, Gould & Bear, 1982; Yumoto, 1988). The amount of in-harmonic spectral components correlates to the perception of hoarseness of the pathological voice (Gelfer, Andrews & Schmidt, 1991). Jotz, Cervantes, Abrahao, Settanni and de Angeli (2002), assessed the efficacy of computerized noise-to-harmonic ratio (NHR) to quantify perceptual and endoscopic findings of dysphonia and or structural lesion of the vocal fold. Their results suggested that NHR is a useful quantitative index to confirm dysphonia, but not the structural lesion. Their notion was that, the risk for having dysphonia increased approximately twice with each increase of 0.01 in NHR. The study thus suggested that, computerized analysis should be used as a complement rather than a substitute for perceptual evaluation.
Voice Turbulence Index (VTI) correlates with the turbulence components caused by incomplete or loose adduction of the vocal folds. VTI is a ratio of the spectral in-harmonic high-frequency energy (2800-5800Hz) to the spectral harmonic energy of (70-4500Hz). This parameter measures the relative energy level of high frequency noise and attempts to compute “breathiness” in the voice signal (Gonzalez & Carpi, 2004).

Soft Phonation Index (SPI) is an evaluation of the poorness of high-frequency harmonic components that may be an indication of loosely adducted vocal folds during phonation. SPI is a ratio of the lower-frequency (70-1600Hz) to the higher-frequency (1600-4500Hz) harmonic energy. SPI provides information related to glottic closure. SPI value and asthenic voice quality are significantly related (Dogan, Midi, Yazici, Kocak, Gunal & Sehitoglu, 2007). Munoz, Mendoza, Fresneda, Carballo and Lopez (2003) and Mathew and Bhat (2009), opined that, SPI is a sensitive parameter to abnormalities in vocal fold adduction.

Bhuta, Patrick and Garnett (2004), studied the correlation between the Grade, Roughness, Breathiness, Aesthenia, Strain in GRBAS scale (a subjective measure of voice) with the Multi-Dimensional Voice Program (MDVP) scale (an objective measure of voice). Of the 19 acoustic variables measured by the MDVP system, the authors observed only three noise parameters significantly correlated with the GRBAS perceptual voice analysis: Grade with Voice turbulence index (VTI), Noise Harmonic Ratio (NHR) and soft phonation index (SPI); roughness with NHR; breathiness with SPI and aesthenia with SPI.

**Tremor analysis**

Tremor is a low-frequency fluctuation in amplitude or frequency (or both). It is natural quality of the human voice provided its extent does not exceed certain limits (Titze, 1994).
The pitch extraction process yields the amplitude and frequency demodulation curves of the voice signal. These curves contain information about the long-term amplitude and frequency variability (tremor) of the voice signal (Winholtz & Ramig, 1992). The principle perceptual features of organic tremor are regular wavering of pitch and intensity, measurable during sustained pitch productions at a range from 4 to 7 Hz. In severe forms, voice breaks may be noticeable (Stemple et al., 2000).

The different acoustic measures which can give information related to the vocal tremors are;

- **Fundamental frequency tremor frequency (Fftr) /Hz/**: This shows the frequency of the most intensive low frequency F0-modulating component in the tremor range.

- **Amplitude Tremor Frequency (Fatr) /Hz/**: This shows the frequency of the most intensive low frequency amplitude-modulating component in the tremor range.

- **Frequency Tremor Intensity Index (FTRI) /%/**: Ratio of the frequency magnitude of the most intensive low-frequency modulating component (F0- tremor) to the total frequency magnitude of the analyzed signal.

- **Amplitude Tremor Intensity Index (ATRI) /%/**: Ratio of the amplitude of the most intensive low-frequency amplitude modulating component (amplitude tremor) to the total amplitude of the analyzed signal.

As per the manual from Kay Elemetrics Corp (1993) the algorithm employed for tremor analysis detects the strongest long-term periodic frequency and amplitude modulations of a signal. The magnitudes of these modulations are quantified by ATRI and FTRI. If the magnitudes of tremor are above a set low detection threshold, Fatr and Fftr provide
quantitative information about the rates of change in tremor frequency and amplitude, respectively.

Shao, Mac Callum, Zhang, Sprecher and Jiang (2010), analyzed the usefulness of acoustic parameters; jitter%, shimmer, Amplitude Tremor Intensity (ATRI), Frequency Tremor Index (FTRI), Amplitude Tremor Frequency (Fatr), and Fundamental Frequency Tremor Frequency (Fftr) in pathological conditions exhibiting vocal tremor (Parkinson’s disease and vocal polyps) in comparison with normal controls. Results indicated significantly higher perturbation measures and tremor parameters in tremulous voice as compared to normal controls. However, significant difference was observed only for frequency tremor parameters than amplitude tremor parameters. Further, the authors found all these parameters having larger standard deviation values, the greatest being for tremor amplitude. Many signals had ATRI and Fatr values of ‘0’, indicating statistical insignificance of these parameters. The authors concluded that, application of these measures required further investigation for the quantification of vocal tremor severity.

**Voice break related measurements**

- Degree of Voice Breaks DVB $/\%/$ - the ratio of the total length of areas representing voice breaks to the time of the complete voiced sample.
- Number of Voice Breaks NVB. The criteria for voice break area can be a missing impulse for the current period or an extreme irregularity of the pitch period (Kay Elemetrics Corp., 1993).
Sub-harmonic components related measurements

- Degree of Sub-Harmonics (DSH) /%/- Relative evaluation of sub-harmonic to F0 components in the analyzed sample.
- Number of Sub-Harmonic Segments NSH - Number of sub-harmonic segments found during analysis

Omori, Kojima, Kakani, Slavit and Blaugrund (1997), analyzed the relationship between acoustic parameters (jitter, shimmer, sub-harmonics) and perceptual rating of roughness in dysphonic cases. The 20 cases that had sub-harmonics were judged as rough voices by perceptual analysis, and jitter and shimmer measures showed five of them having normal jitter and seven of them having normal shimmer values. These results suggest that, perception of roughness was better correlated with presence of sub-harmonics than the jitter and shimmer parameters. Thus, it can be presumed that, sub-harmonic analysis in the power spectrum provides an objective parameter for evaluating rough voice.

Voice irregularity related measurements

- Degree of Irregular Vocalization (DUV) /%/- the number of unvoiced segments detected during analysis.
- Number of Unvoiced Segments NUV- relative areas of non-harmonic areas in the voice sample.

Very limited studies have been conducted on the acoustic characteristics of call center operator’s voices. Lehto et al., (2006); (2008), studied the correlation between subjective voice complaints and acoustic measures in 24 female Customer Service Advisors, during one working day, at four different times. Subjective voice complaints were studied using self-
reported questionnaires and the acoustic measures included; fundamental frequency (F0), Sound pressure level (SPL), Alpha Ratio (AR) and the number of vocal fold vibrations (Index= F0 x phonation time). To represent the vocal symptoms, they used three variables; vocal fatigue, hoarseness and general sum variable (12 statements on vocal symptoms). Their results revealed that, time of the day did not have statistical significant effect on variables like SPL, AR and Index except for F0 which rose significantly during working day. In subjective measures, time of the day showed a very large effect on all three subjective measures (fatigue, hoarseness and sum of variable). The researchers, however, did not find any significant linear correlation between the subjective and objective data.

Shekhawat, Kant, Shah and Dhola (2006) compared voice of call centre professionals working in call centers for more than 6 months and less than a year with those working for more than 1 year and controls (those not employed in call centers and also having no voice complaints) using acoustic and perceptual analysis. They reported a significant difference in jitter, shimmer, number of unvoiced segments, and fundamental frequency between the call center workers and controls. However, they did not find any difference within the call center operators (working less than a year and those working for more than a year). On perceptual analysis, they found the voice quality (call center operators) to be mildly hoarse in 26%, mildly harsh in 7%, mildly breathy in 7%, moderately hoarse in 6%, and normal in 54%.

Shah and Sanghi (2010), compared the acoustic parameters such as F0, perturbation measures (jitter and shimmer) and Normalized Noise Energy (NNE) between International Call Centre employees (ICC), Domestic Call Centre employees (DCC) and Non Vocal Professionals (NVP). Among the three groups, significant differences were observed in
amplitude perturbation measures and NNE measures. There was no significant difference between ICC and DCC groups in terms of amplitude perturbation measures. However, the NNE values significantly varied between ICC and DCC, ICC and NVP, and DCC and NVP groups.

Some of the studies in the literature indicated that acoustic measures are sensitive to detect the presence of dysphonia. Munoz et al., (2003), studied the acoustic correlates to vocal quality in group of men and women with and without voice disorders, based on evaluation of a group of judges experienced in the field of vocal rehabilitation. In male subjects, perceptual evaluation of normal, hoarse and rough voice qualities were related to the following acoustic features: frequency perturbation measures (JITA, RAP, and SPPQ), amplitude perturbation (SAPQ and VAM), soft phonation index (SPI) and fundamental frequency tremor intensity (FTRI). However, quality wise, female voices were perceived as normal, breathy and hoarse, but the acoustic correlates were less conclusive.

Ma and Yiu (2006) evaluated the accuracy of acoustic voice measures (F0, Jitter, shimmer and NHR) in predicting the perceived overall severity of voice quality. The dysphonic groups demonstrated significantly lower mean fundamental frequency values and significantly higher jitter and shimmer values than the control group. However, the noise-to-harmonic ratio values of both groups were similar and were not significantly different. According to these authors, perturbation measures (jitter & shimmer) are sensitive enough to detect the presence of dysphonia and enhance the clinical value of dysphonia evaluation.

Niebudek-Bogusz, Kotylo and Sliwinska-Kowalska (2007), studied the possible acoustic signs of vocal fatigue (comparing pre and post loading) in 51 female teachers with
functional voice disorders. Acoustic analysis after vocal loading revealed an increase rate of
abnormal pitch perturbation quotient (PPQ), jitter, relative average perturbation (RAP) as well
as amplitude perturbation measures. These authors suggested that, acoustic measures can be
used to identify early stages of dysphonia in professional voice users.

Laukkanen, Ilomäki, Leppänen and Vilkman (2008), investigated the relationship
between the symptoms of vocal fatigue to acoustic variables reflecting type of voice
production (habitual loudness and loud). Acoustically, the participants exhibited higher F0,
SPL, and alpha ratio and subjectively, tiredness as a symptom of voice after a working day.
Increase in jitter and mean F0 in loud reading correlated with tiredness of throat. The authors
suggested that, differences in the acoustic parameters after a vocally loading working day
mainly seem to reflect the increased muscle activity as a consequence of vocal loading.

However, not all the studies in the literature find acoustic parameters having
significant association with vocal symptoms. Eustace, Stemple and Lee (1996) measured
acoustic, aerodynamic and video-stroboscopic data on 88 patients with a primary complaint of
vocal fatigue. Video-stroboscopic evaluation revealed presence of anterior chinks, anterior
and posterior chinks, and spindle shaped vibratory closure patterns. Aerodynamic measures
were characterized by elevated airflow rates, reduced maximum phonation time indicating
hypoadduction of the vocal folds. However, acoustic measures such as F0, F0 range and jitter
were within normal limits. The results of this study show that, even though acoustic measures
are claimed to differentiate between normal and pathological voices, it is not clear whether,
they represent pathological or deviant voice quality consistently and reliably. This finds
support in the study by Giovanni, Robert, Estublier, Teston, Zanaret and Cannoni (1996) who
reported that, acoustic and aerodynamic measures cannot discriminate mildly impaired voice quality. Gelfer (1999) states that, voice is a multidimensional phenomenon, comprised of a number of elements that contribute to overall voice quality and vocal effectiveness. Thus, these elements often interact in ways that are difficult to measure objectively, which focus on discrete and isolated voice components by their nature. Most objective measures are based on isolated vowel samples and measures like jitter and shimmer are based on the assumption to maintain steady vocal frequency and noise to harmonic ratio to maintain a stable formant structure. Thus, according to these authors, even though the acoustic measures are said to differentiate between normal and pathological voices, they cannot measure the salient attributes that listeners use in judging the voice.

**Auditory perceptual assessment findings in professional voice users**

Auditory perceptual evaluation is the most commonly used clinical assessment method to identify the disordered voice quality (Behrman, 2005). One definition of auditory perception is: "Identification, interpretation, or organization of sensory data received through the ear" (Nicolsi, Harryman & Kresheck, 1978). Fex (1992) says: "The term perceptual evaluation of a person's voice means that a listener is making a comparison between a not necessarily specified 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". Definition of normal voice is not certain as a vast number of people are supposed to have normal but individually different voices. These definitions indicate that normal voice quality is based on subjective opinion and may vary with different cultures. It is considered to be basic foundation of voice evaluation and treatment, and the comparative standard of this evaluation will assist the clinician to relate and interpret acoustic and physiological measures.
with this. Perceptual measures provide baseline information about the degree and nature of clients’ voice problems. Perceptual description of voice qualities are easily interpreted by a wide range of people than descriptions of the instrumental measurements (Oates, 2009). Gelfer (1999) opines that, perceptual measures can provide a global picture of the patient’s vocal performance. These measures can furnish significant information on how various vocal attributes may be affected by linguistic factors, prosody and other variables. Perceptual evaluation will also consider the individuals vocational and social needs. Thus, voice is a multidimensional phenomenon, comprised of a number of elements that contribute to overall voice quality and vocal effectiveness. The complex interactions of these elements are difficult to measure objectively.

The term quality refers to characteristic of sound that distinguishes from other sounds of similar pitch, loudness and duration. In other words, voice quality is the auditory perception of acoustic elements of phonation that characterize an individual speaker. Thus, it is an interaction between the acoustic speech signal and a listener’s perception of that signal. Various research methods provide insight as to what is meant by normal voice quality. Eskanazi, Childers and Hicks (1990) consider normal voice “as a voice with no apparent pathology, and no unusual voice characteristics or habits”. Karnell et al., (1991), described normal voice quality as having no hoarseness.

Even though the definition of normal voice is not very consistent, listeners have similar, relatively stable standards for ‘normal’ voice quality (Kreiman, Gerratt, Precoda & Berke, 1992). This could be attributed to the everyday experience the listeners have with normal voices. Hence, consistency is observed when they rate normal or near normal voices.
Similarly, the internal standard for pathological voice may vary from a listener to listener depending upon their experience or exposure to it. It’s suggested that, listeners need many years to develop a stable set of criteria for rating pathological voices reliably as pathological voices are characterized by great acoustic variability (pitch, loudness, spectral envelop and its changes over time, any noise components, fluctuations in F0 and other factors) within and across both utterances and speakers (Gerratt & Kreiman, 2000). Thus, pathological voice qualities associated with complex psychophysical functions are difficult to specify and bring greater disagreement in the ratings of voice quality.

Supporting these views, Niimi and Miyaji (2000), reported that, various vocal fold pathologies can produce an identical voice quality or different voice quality in two cases having similar diagnosis. These authors analyzed the vocal fold vibration patterns and associated pathological voice quality using high speed digital imaging systems in 22 patients with vocal fold lesions. They observed three types of vocal vibratory patterns; a group with complete glottal closure, a group with incomplete glottal closure, and a group with inconsistent glottal closure. These patterns of glottis closures had significant effect on perception of vocal quality (overall severity/grade). The voice rated to have high scores on grade and roughness showed asymmetric or irregular vibration or vibrations with different frequencies between the vocal folds. The perceived breathy quality was associated with incomplete closure, inconsistent closure or asymmetry of amplitude and mucosal wave. This study indicates that, a type of pathological voice could be produced by a particular vibration of the vocal fold and not by pathology of vocal fold.
Rabinov, Kreiman, Gerratt and Bielamowicz (1995), reported that, listener perception and acoustic analysis packages differ greatly in their measurement characteristics of pathologic voice. The authors studied the relative reliability of human listeners and automated systems for measuring perturbation in the pathologic voices. They concluded that, auditory perceptual ratings by the listeners get more reliable when the severity of the vocal pathology increases, whereas, objective methods quickly broke down as severity increased. They also found that, human listeners’ rating of roughness was much consistent than objective algorithms.

Review of literature reveals several quality dimensions that have been consistently used to categorize dysphonic voice quality.

**Breathy voice**

It is commonly perceived as whispery or airy voice. In a physiological view, this term is associated with hypo-adduction (i.e., lack of vocal fold adduction or closure) of the vocal folds and refers to the audible detection of airflow through the glottis during phonation. The combination of audible airflow with quasiperiodic voice signal results in airy or hazy characteristic of the breathy voice (Awan, 2001). The presence of any disease that results in the formation of a glottic chink during phonation can produce a breathy voice (Fukazawa, Blaugrund, EI-Assuooyt & Gould, 1988). Even though this quality of voice is considered as a common attribute of disordered voice, a slight amount of breathy voice can be a common feature of normal female voice (Klatt & Klatt, 1990). Perception of breathy voice quality can be related to the changes in the partial loudness of the harmonic energy with the aspiration noise acting as a masker (Shrivastav, 2003).
Harshness/Roughness

The harsh voice is usually perceived as unpleasant voice and associated with terms such as coarse, strident, low pitched noise, and rasping (Askenfelt & Hammarberg, 1986; Gelfer, 1988). In terms of physiology, harsh voice is associated with hyper-adduction (i.e., excessive vocal fold adduction or closure) of the vocal folds and refers to the noise production as a result of irregular vocal fold vibration (Askenfelt & Hammarberg, 1986). Roughness of the voice can be caused by a soft swelling or a mass imbalance of the vocal cord. Jitter and shimmer are the two acoustic parameters strongly associated with the roughness (Fukazawa et al., 1988).

The term roughness is synonymously used with harshness (Wolfe, Fitch, & Martin, 1997). Even though both the terms refer to irregularity of vocal fold vibration, the rough voice may be characterized by the presence of low frequency noise components versus a high frequency noise found in harsh voice (Askenfelt & Hammarberg, 1986). Current terminology in the perceptual assessment of voice appears to favour roughness over the term harshness (Wolfe et al., 1997).

The perception of roughness can result from interference of two simple sinusoidal tones. The degree of interference among the harmonics of two sinusoids determines the amount of dissonance or consonance perceived for simple frequency ratios of the fundamentals. The roughness can also result from amplitude or frequency modulation (Bergan & Titze, 2001).
Hoarseness

Hoarseness can be the result of the increased mass that results from a vocal nodule or polyp or localized edema. Sometimes, it can be observed in the absence of a mass lesion of any sort (Johnson, 1993). The term hoarseness is a hybrid descriptor that denotes a voice with both breathy and harsh/rough qualities simultaneously. Therefore, hoarseness originates from a combination of irregularity of vocal fold vibration and turbulent airflow through the glottis (Omori et al., 1997). This voice quality arises when an individual attempts to compensate for a breathy voice by using increased laryngeal tension and hyper-adduction resulting in a harsh or rough component. Many authors have concluded that hoarseness is the most common voice quality characteristic (Andrews, 1995; Boone & McFarlane, 1988).

Strain

This is defined as the impression of vocal effort, i.e., a hyper-functional state of phonation (Hirano, 1981; deKrom, 1994). The effortful characteristics are consistent with both hypo-adductive and hyper adductive states in which the speaker must use increased muscular force and tension in the attempt to produce phonation, resulting in strident harsh vocal quality. It is a most unpleasant voice form from an aesthetic perspective. This type of voice pattern is associated with musculoskeletal tension disorders. The patient with strained voice quality complains of discomfort while speaking, with obvious physical strain. The “tightness” in the strained voice is accomplished through tension in the larynx, as well as supra-laryngeal vocal tract (Johnson, 1993). Thus, the perceived vocal quality is best understood as a combined disorder involving both voice and resonance.
Roughness and breathiness features are associated with organic lesions in which there is a lowering vibration (roughness) and default of closure (breathiness), whereas features of asthenia and strain are associated with functional voice disorder, related with vocal tiredness and hyper-functional emission (Sáenz-Lechón, Godino-Llorente, Osma-Ruiz, Blanco-Velasco, & Cruz-Roldán, 2006). There is a great variation in types of voice distortion that can be caused by a disorder, ranging from perceived breathy voice quality to strained strangled voice quality. There is no single parameter that can distinguish a normal from an abnormal voice, particularly when the alterations to the voice are only mildly or moderately distorted (Sapienza and Woodson, 2009).

**Voice samples used in the Perceptual Assessment of Voice Quality**

A major emphasis in the perceptual evaluation of voice disorders is placed on the analysis of sustained vowels, quite often the vowel /a/ (deKrom, 1994). However, many investigators have concluded that sustained vowels do not adequately represent continuous speech. For example, according to Askenfelt and Hammarberg (1986), a sustained vowel tends to be representative of voice function status only in those cases where the voice is affected by a severe laryngeal pathology, such as laryngeal cancer or unilateral paralysis. Consequently, for the majority of the subjects, it is necessary to analyze running speech in order to obtain an adequate estimation of the voice status. Furthermore, according to Takahashi and Koiki (1975), the initial and the terminal parts of the voice may carry abundant information not contained in the steady-state vowel. On a similar note, Hammarberg, Fritzell, Gauffin, Sundberg and Wedin (1980) stated that "Changes in running speech such as vocal onset and termination, voice breaks, etc., are crucial to voice quality, and are not likely to appear in a single vowel sound".
Taking these points into consideration, several studies have been conducted to assess the effect of stimulus type on the perception of voice quality. deKrom (1994), examined the consistency and reliability of perceptual voice quality ratings from sustained vowels versus continuous speech samples, and reported that reliability of ratings on connected speech stimuli was lower than vowel segments. The possible reasons quoted for such variations were distraction of the listener from non-voice source information (e.g., dialect, rate) and or variability of voice quality within the connected speech segment. He opined that vowel segments appeared to be more acceptable for both acoustic and perceptual assessment. However, studies conducted over the recent years did not agree on this concept. Wolfe, Cornell and Fitch (1995), evaluated the perceptual relationships between sustained vowels and connected speech samples in a group of commonly occurring laryngeal conditions and a group of normal speakers. They observed that, the vowels were judged to be more severe than sentences when all groups were combined (laryngeal pathology and normal speakers). When this effect was studied in individual groups, no significant differences were observed in rating the severity of voice using vowel and sentences for pathologic groups. With respect to normal voice group, the vowels were judged to be more severe than the sentences. It was hence concluded that, normal speakers frequently produced sustained vowels that are more dysphonic than continuous speech which they attributed to the fact that sustained vowels were produced under unnatural conditions and would have probably caused this perception of dysphonia among the normal speakers.

Revis, Giovanni, Wuyts and Triglia (1999), evaluated the effects of stimulus on the perceptual rating voice quality. They included three different voice samples i.e., connected speech (P), a sustained vowel including both initial and stable portion (complete sustained
vowel; CSV), and a sustained vowel including only the stable portion (SSV) from 60 dysphonic and 20 normal controls. Their results indicated that, stabilized sustained vowels were judged less severe than judgments on connected speech and good correlation between the ratings of CSV and connected speech. This was however, not true for perceptual attribute of breathiness, i.e., a good correlation between SSV and connected speech (0.926) than CSV and connected speech (0.653) was observed. This made them state that, perceptual analysis using only the stable portion of vowel underestimated dysphonia. However, it was not true for all the perceptual parameters, as there was upper estimation of breathiness during CSV sample. Thus, sustained vowels cannot be considered as pertinent for perceptual voice analysis as subjects may change their style of voice production while sustaining vowels.

Zraick, Wendel and Smith-Olinde (2005), investigated the effects of speaking task on auditory perceptual judgment of the severity of dysphonia using three tasks; sustained vowel /a/, oral reading of a standard passage, and connected speech describing standard picture. Their study included voice samples of these three tasks from 29 dysphonics. A statistically significant effect of speaking task on the auditory perceptual analysis of voice quality emerged. Across the speakers, listeners rated the dysphonia as more severe when they heard the sustained vowel than the other speaking tasks. No significant difference was observed between oral reading and picture description.

These three samples required the subject to produce voice under circumstances of differing vocal function and give the clinician a standardized method that could be repeated from one subject to the other, as well as allowing comparisons of voice quality deviations. Recent studies suggest that perceptual assessment of voice should include sustained vowel
segments, standardize sentence/passage reading and continuous spontaneous speech samples (Awan, 2001).

**Perceptual voice assessment scales**

Basically, perceptual voice rating scales are classified ordinal scale and visual analog scales. The ordinal scale rates the severity of voice problem using four or seven levels of severity. On a four point scale, voice quality is graded as normal (grade 0); slight dysphonia (grade 1); moderate dysphonia (grade 2) and severe dysphonia (grade 3). The seven point scale includes three extra ‘half tone’ levels i.e., 0.5 for normal to slight, 1.5 for slight to moderate, and 2.5 for moderate to severe (Yu, Revis, Wuyts, Zanaret & Giovanni, 2002). Due to its simplicity, the ordinal scales are most widely used in perceptual analysis of voice. However, these scales do not allow raters to give full expression to their listening acuity (Gerratt et al., 1993; Karnell Melton, Childes, Coleman, Dailey, & Hoffman, 2007). Visual analog scales consist of 100 mm long line with left end representing maximum severity. The listener has to indicate the perceived degree of severity by making a mark at a distance corresponding to the perceived degree of severity. The score is determined by measuring the distance in millimeters from the left hand end and the mark (Yu et al., 2002). Visual analog scale provides better discrimination of degree of severity compared to ordinal scales (Kreiman, Gerratt, Kempster, Erman & Berke, 1993; Karnell et al., 2007; Kempster, Gerratt, Abbott, Barkmeier-Kraemer, & Hillman, 2009). According to Yu et al., (2002), visual analog scale is difficult to interpret as they increase the variability and there are no predictable landmarks to interpret the degree of severity. To overcome this limitation, the authors suggested transformation of the visual analog scale into a four point scale with the help of weighted conversion scheme. Taking this into consideration, they developed the modified visual analog scale.
scale (mVA). This choice was based on the assumption that normal voices and severe
dysphonia were more easily identified by the judges and that weighted conversion scheme
should have more freedom in quantifying intermediate dysphonia in the middle of the scale
than normal voices and severe dysphonia at the extremities of the scale. It was also suggest
that, the VA scale cannot be divided in to equal scoring ranges to obtain mVA scale. To
obtain the mVA scale, mean values are calculated using the ten (10 cm) VA scores
transforming into four grades according to the weighted conversion scheme as follows;

<table>
<thead>
<tr>
<th>VA scale cm</th>
<th>mVA scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 0.9</td>
<td>0 (G0)</td>
</tr>
<tr>
<td>1.0 – 5.0</td>
<td>1 (G1)</td>
</tr>
<tr>
<td>5.1 – 9.0</td>
<td>2(G2)</td>
</tr>
<tr>
<td>9.1 - 10</td>
<td>3(G3)</td>
</tr>
</tbody>
</table>

Using the mVA scale, Yu et al., (2002), analyzed the variability of perceptual
judgment as a degree of severity and found greater variability in G1 and G2 than G0 and G3.
They opined that, greater variability in the judgment was probably due to poor internal
standards to distinguish intermediate grades of dysphonia. Thus, neither the ordinal nor the
VA scales fully meet the challenge of classifying intermediate dysphonia. However, these
authors found that, mVA are better than conventional ordinal scale for the discrimination of
intermediate grades (G1 and G2).

There are different perceptual voice assessment scales to describe the quality of an
individual’s voice. For e.g., GRBAS scale, Buffalo Voice Profile, Vocal Profile Analysis
Scheme, CAPE-V etc. Among them, GRBAS scale of the Japan Society of Logopaedics and
Phoniatics is in most widespread use. Parameter ‘G’ in GRBAS is quantitative in nature and
reflects the overall impression of abnormality in voice. Roughness psycho-acoustically refers
to irregular perturbation of pitch and amplitude and noise in low frequency region. The physiologic findings associated with these psychoacoustic differences include soft and swollen vocal folds and asymmetry of vocal fold mass. Breathiness scale is psychoacoustically characterized by noise below the mid-frequencies and physiologically, incomplete closure of vocal folds leading to high expiratory airflow. Asthenic scale is associated with less harmonic content in the high frequency region, irregularity of pitch and amplitude and a fading amplitude contour (psycho-acoustically) and vocal hypo-function physiologically. The strain scale reflects characteristics such as a higher pitch, noise in the higher frequencies, increased amplitude of the higher harmonics, and increased pitch and amplitude perturbation (psychoacoustic). Physiologically, this scale is sensitive to vocal hyper-function, stiff vocal folds, and increased vocal fold mass (Yamaguchi, Srivastav, & Andrews, 2003; Saenz-Lechon et al., 2006).

Even though different perceptual voice assessment scales are available, literature has failed to offer any recommendations concerning the best scale in regard to robustness, consistency and inter-rater agreement. Moreover, these perceptual analysis scales do not specify the use of clinical administration of speech material or standardized protocol to be followed. Even though GRBAS scale has been widely used for judging voice quality, it does not offer a specific protocol for administration and does not provide guidelines for analysis (Kempster et al., 2009). Looking into these limitations, a consensus meeting sponsored by American Speech-Language-Hearing Association’s (ASHA) Division 3: Voice and Voice disorders, and the Department of Communication Science and Disorders, University of Pittsburgh, resulted in the development of an auditory perceptual instrument named the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) in 2002. It represents the
consensus committee’s minimum recommended standard for the auditory perceptual analysis of voice disorders and is intended to facilitate increased consistency across raters.

**CAPE -V**

The Consensus Auditory Perceptual Evaluation of Voice (CAPE-V) is a recent perceptual voice analysis instrument. It includes a specific protocol that designates the tasks (phonation, reading and spontaneous speech), procedures and scaling routine, towards the larger goal of improving the consistency of clinical assessment without excessive demands on the clinicians’ time or learning. It has included a 100 mm visual analog scale allowing to measure the voice quality in the millimeter interval to better accommodate the multidimensional features of voice quality. The CAPE-V has included six salient perceptual vocal attributes, identified by the core consensus group as commonly used and easily understood. The attributes are: (a) Overall Severity (Global, integrated impression of voice deviance); (b) Roughness (Perceived irregularity in the voicing source); (c) Breathiness (Audible air escape in the voice); (d) Strain (Perception of excessive vocal effort); (e) Pitch (Perceptual correlate of fundamental frequency, relative deviation from normal for the person’s age, gender etc); and (f) Loudness (Perceptual correlate of sound intensity, relative deviation from normal for person’s age, gender etc). Along with this, CAPE- V also attempts to document more voice quality features, allowing for supplemental feature scales and comment areas. For eg; if the subject exhibits resonance problem, it has to be indicated as additional feature. Additional scoring will not be made for this. For each dimension, scalar extremes are unlabeled. The scale has only general regions indicated as “MI” (mild deviancy), “MO” (moderately deviant) and “SE” (severely deviant). For each attribute, listener can comment on the presence of consistency (C) or inconsistency (I) across or within the tasks.
The judgment of consistency indicates the presence of that attribute continuously throughout the tasks. A judgment of “intermittent” indicates that the attribute occurred inconsistently within or across tasks. For example, an individual may consistently exhibit a strained voice quality across all tasks, which include sustained vowels and speech. In this case, the rater would circle “C” to the right of the strain scale. In contrast, another individual might exhibit consistent strain during vowel production, but intermittent strain during one or more connected speech task. In this case, the rater would circle “I” to the right of the strain scale. For one subject, one rating scale will be used and if there is difference in the vocal attributes across the tasks, it will be indicated using task numbers on the rating scale (Karnell et al., 2007; Kempster et al., 2009).

Karnell et al., (2007), published a preliminary report comparing the reliability of the clinician based auditory perceptual judgment using CAPE-V and GRBAS voice rating scales. They found comparable estimates of inter-rater reliability for both the scales. However, they suggested that the CAPE-V may offer more sensitivity to small differences within and among patients than the GRBAS scale. Clinician’s perceptions of dysphonia appeared to be reliable and unaffected by rating tool, as indicated by high level of agreement between the two rating systems when they were used together. Similarly, Kelchner, Berhm, Weinrich, Middendorf, de Alarcon, Levin and Elluru (2008), examined the reliability of perceptual voice analysis in pediatric voice disorders using CAPE-V. The authors reported excellent agreement within and across three raters from the same settings has been reported.

Zraick, Kempster, Connor, Thibeault, Klaben, Bursac, Thrush, and Glaze (2010), assessed the reliability and validity of CAPE-V in judging the normal and disordered voices.
For this purpose, they examined the intra and inter rater reliability of experienced voice clinicians’ judgments of voice quality using the CAPE-V and GRBAS and assessed the relationship between the two scales to establish the validity of CAPE-V. They found four scales of GRBAS and CAPE-V were most comparable: overall severity and Grade; roughness and rough; breathiness and breathy; strain and strained. However, the results indicated slightly higher inter-rater reliability values for CAPE-V than GRBAS among these four scales. To establish the empirical validity of the CAPE-V, they compared experienced raters’ judgments of voice quality to the judgments made by GRBAS. The results indicated a strong multiserial correlations between the two scales suggesting empirical validity of CAPE-V.

Rabinov et al., (1995), asked experienced listeners to provide ratings of roughness using a visual analog scale (VA) that ranged from normal to severely disordered. They found that listeners could accurately and reliably identify roughness and reliability improved with increase in the severity of pathology. On the contrary, Wuyts, de Bodt and Van de Heyning (1999), reported perceptual evaluation of 14 pathological voices performed by 29 listeners using two versions (ordinal scale and visual analog scale) of GRBAS scale. They found higher agreement with ordinal scale compared to the visual analog scale. According to them, visual analog scales provided increased freedom of judgment and decreased the inter-rater agreement.

Regardless of the critical problems that plague perceptual judgments of voice, this analysis modality holds clear face validity and functional importance. It allows complex and sophisticated judgments of quality far superior in many ways to the algorithms that underlie acoustic indices.
Reliability of perceptual evaluation

Evaluation of voice quality is an important part of the diagnosis and treatment of voice disorders. Despite this substantial importance of voice quality, measurement of voice quality is found to be problematic, both clinically and experimentally. There are ongoing researches to better understand and account for the unreliability of auditory perceptual voice quality ratings. Certain variables such as type and complexity of the rating task, listener’s background, training, type of voice samples are reported to affect perceptual judgment (Oates, 2009). Kreiman and Gerratt (1996) reported of disagreement between researchers in using a standardized set of scales for assessing voice quality due to large differences between listeners’ perceptual strategies. Despite the perceptual evaluation of voice considered as a gold standard of voice evaluation, reliability (intra-rater and inter-rater) of perceptual evaluation is found to be poor.

The poor reliability of perceptual evaluation of voice is mainly related to poorly defined perceptual voice characteristics, severity of the perceived voice disturbance, experience of the listener and lack of stable internal standards of the listeners. As the listeners gain experience, their internal representation and perceived assessment of severity for a dysphonic voice changes. This can lead to different perceptual strategies in evaluating dysphonic from normal voices (Kreiman et al., 1993). The authors further stated that, reliability of perceptual assessment of voice was good at perceptual extremes, that are normal Vs. severe distinction, whereas the perceptual categories in close proximity such as normal Vs. mild and moderate Vs. severe were more difficult to differentiate reliably. Perceptual qualities (breathiness, roughness, harshness, hoarseness etc.,) have never received widely accepted definitions in the clinical literature. Thus, it is difficult to determine precisely what a
particular author or listener means by “hoarseness”, “breathiness”, “roughness”, or any other label for vocal quality. Thus, listeners often disagree when they rate vocal qualities, as there is significant individual differences exist in the meaning assigned to such terms (Kreiman, Gerratt & Berke, 1994). These authors further examined whether, judgment of breathiness and roughness were independent of each other. Authors reported that, roughness did not appear to influence raters’ judgments of breathiness, but judgment of roughness were heavily influenced by degree of breathiness. The findings indicated, different aspects of a quality are apparently a significant source of inter-rater unreliability in ratings of pathological voices. According to the reports of Yiu, Murdoch, Hird and Lau (2008), cultural and language background of listeners also has an influence on the perceptual voice rating. Kreiman et al., (1993), found that, sequence of sample presentation has an important role in variability. That is, a voice sample classified as intermediate dysphonia, may be overrated or underrated depending on whether it is presented after a normal or severely dysphonic voice sample. It has hence been suggested to avoid such ‘sequence effect’ by presenting the voice samples three times in random order during three different listening sessions.

DeBodt, Wuyts, Van de Heyning and Croux (1997) explored the effect of experience and professional background on perceptual rating of voice quality using the GRBAS perceptual voice analysis scale. They presented nine pathological voices to speech language pathologists (experienced) and otolaryngologists (inexperienced) with a 14-day test-retest interval. The test-retest reliability was found to be moderate, with best agreement occurring for the G (grade) parameter and the worst for the S (strained) parameter. The study indicated that the reliability of perceptual voice evaluation depends on perceptual attributes measured than the level of experience or professional background.
Kreiman, Gerratt and Ito (2007), identified four experimental design factors affecting the perceptual analysis of voice; internal memory standards for levels of perceptual dimension, ability of the listener to isolate single dimension in a complex context, resolution of perceptual analysis scales, and absolute magnitude of the attribute being measured. Shrivastav, Sapienza and Nandur (2005), have shown that listener variability could be minimized by applying psychometric principles when designing the listening task. Such principles include averaging the ratings from multiple presentations of the same stimulus, with the investigators showing that a minimum of five repetitions provide the best results for judging voice quality in sustained vowels. In addition, to account for scale resolution and edge effects related to the absolute magnitude of a perceptual attribute, a standardized value for each rating is computed. The averaging and standardization procedures attempt to minimize variability and response biases of individual listeners. Although this approach allows for the experimenter to obtain more reliable perceptual ratings of natural (versus synthesized) vowel stimuli, this has limited clinical application due to the impracticality of taking the time to have several clinicians independently rate multiple repetitions of the same voice samples.

Aakanksha (2009), studied the reliability of perceptual evaluation (21 dysphonic voices) of voice using consensus CAPE-V across different tasks (phonation, sentence reading, spontaneous speech) given in the CAPE-V scale. The reliability of perceptual ratings of dysphonic voices was computed using Cronbach’s alpha reliability coefficient. The results indicated high reliability for perceptual evaluation of dysphonic voice across the tasks and judges except for loudness which was reported as moderate. The study also examined the most appropriate task in the CAPE-V for perceptual evaluation of voice. Spontaneous speech
task was having highest inter-rater reliability \( r = 0.81 \), than oral –nasal sentences \( r = 0.73 \) and phonation \( r = 0.70 \). However, the study findings by Wuyts et al., (1999), pointed out that, though visual analog scale offers finer judgment of voice quality, with increased freedom, the inter-rater agreement decreases considerably. These authors also reported that, the listeners have the general tendency of rating at the middle of the 100 mm line and opined that, visual analog scale has more variability in rating the voices than the ordinal scale.

### 2.8 Occupational voice care

In the management and treatment of voice disorders, importance is placed on education and explanation of vocal tract care and vocal hygiene (Mathieson, 2001). Voice care knowledge includes an understanding of issues of vocal hygiene, limiting vocal hyper-function, and reducing vocally abusive behaviours (Pannbacker, 1998). The goal of voice care is to restore the best voice possible; a voice that will be functional for the purpose of employment and general communication (Colton, Casper & Leonard, 2006). One of the greatest risk factors for professional speakers is lack of professional vocal training (Jacobson, 1993). These professionals should hence be cautioned about the factors which assist them in maintaining long term success in their vocal health. They should be educated to identify factors that may contribute to voice problems, to alter and avoid them and modify vocal behaviour before any damage occurs (Duffy, 2003).

Call center employers often fail to take adequate steps to protect their employees’ health and safety. As a white-collar workplace, the potential hazards within the workplace of CCO are not outwardly obvious. The risks “tend to involve cumulative stresses and strains resulting from a combination of causes rather than a single traumatic event” making it
impossible to pinpoint any one, or rather, any precise combination of factors likely to have caused the harm (Paul & Huws, 2002). Due to the high prevalence of vocal attrition in telemarketers, recent studies have emphasized the necessity of vocal training programs for these high risk professionals with their requirement of efficient voices for telemarketing operations (Jones et al., 2002; Lehto et al., 2003; Taylor & Oates, 2004). Studies in the literature, based on subjective judgments, have shown that, CCOs are getting some benefit from vocal training (Lehto et al., 2003, Lehto et al., 2005; Olivera et al., 2009). Lehto et al., (2003) studied the effects of short vocal training (2 days) on 38 female and 10 male CCOs. This short term vocal training course was found to have significant effect on some of the vocal symptoms (diminished sensation of mucus, throat clearing and worsening of voice) as reported by call centers operators. Overall, more than 60% of the CCOs (male and female) reported that, voice training had improved their vocal habits. Since the effectiveness of the voice training was measured subjectively, these authors suggested more objective measures.

Similar findings have been reported by Lehto et al., (2005) by investigating both the short term and long term effects of vocal training in CCOs. They observed that female subjects gained long term benefits from training than males. However, even though overall level of the voice symptoms was lower after training, there was also an increase in some of the symptoms in the follow up study. This justified the need for regular consultation and training for the occupational voice users to sustain positive effects of voice training. Thus, this study does not allow for any strong conclusions on benefits of vocal training in CCOs.

Oliveira et al., (2009) evaluated the effectiveness of vocal training program in the prevention of vocal attrition in telemarketers. Their voice training module included warm-up,
warm up and cool down, warm up and group activities regarding psychodynamics of voice production. Their vocal training program was divided into eight 30-min modules offered weekly. They measured the effectiveness of vocal training by comparing the vocal symptoms pre and post training as well as using a questionnaire ‘Benefits Obtained with Voice Training’. Results of the study showed that, telemarketers benefited with vocal training and reported improvement in vocal and communication performance, but did not induce changes in vocal symptoms. They concluded that, vocal training may not give effective results in reducing the vocal symptoms; rather, the vocal training may help the telemarketers in preventing the development of the voice problems.

Few attempts have been made to examine the general knowledge base of voice and voice care in CCOs. This could highlight the need for providing preventative education or health promotion to the wider population. On an individual basis, documenting existing voice care knowledge provides a baseline for interventions and is a starting point for designing appropriately targeted treatment programs. Changes in voice characteristics and effects of voice problems in communication are important for the prevention of the occurrence of voice problems in professional voice users (Simberg et al., 2000). Thus, voice education; either with indirect vocal hygiene education or direct exercises and voice training (focused on, e.g., posture, respiration, tension release in the vocal apparatus, resonance and voice projection) are important for dealing with voice problems (Chan, 1994). Early identification and treatment of voice problems is thought to reduce the severity and the time needed to recover from the voice problem for the professional speakers. However, there is a lack of systematic research to show whether or not the voice improves with vocal education, with a direct or
indirect, primary or secondary approach. And along with this deficiency, there is also a lack of systematic voice education in the professional voice users.

Professionals developing preventive voice care programs should understand the vocal tasks expected of the voice user and evaluate their voices at their working situations, and identify those factors contributing to developing dysphonia. Accurate identification of risk factors will enable the laryngologist and speech language pathologist to design a program of behavioural therapy based on the specific nature of the observed pathophysiology. In occupational voice users, the most common cause of voice problems is vocal hyper-function or strain. Early identification and treatment of voice problems is thought to reduce the severity and the time needed to recover from the voice problem for the professional speakers.

Thus, the literature review has tried to cover relevant information and studies related to the analysis of the voice and its problems in CCOs as quoted in the western literature and the limited studies in the Indian context. It is hence an overview of the general vocal disorders encountered in professional speakers with specific attention given to CCO’s.