A Comparison of Trained and Untrained Vocalists on the Dysphonia Severity Index

*Shaheen N. Awan and †Anysia J. Ensslen, *Bloomsburg, Pennsylvania, †Lexington, Kentucky

Summary: The purposes of this study were (1) to compare trained and untrained singers on the Dysphonia Severity Index (DSI) and its component measures, and (2) to contribute to normative DSI data for trained singers. This study included 36 untrained participants (15 males and 21 females) and 30 participants (15 males and 15 females) with singing experience between the ages of 18 and 30 years. Measures of maximum phonation time (MPT), highest phonational frequency, lowest intensity, and jitter were obtained for each subject and incorporated into the previously published multivariate DSI formula. Results indicated that vocally trained subjects have significantly higher DSI scores than untrained subjects (mean DSI: 6.48 vs 4.00, respectively), with significant differences observed between trained and untrained groups for three of the four components of the DSI (F_0 high; I low; jitter). The findings of this study are consistent with previous reports that indicate significant increases in the DSI with vocal training, and with various studies that have observed increased vocal capability in trained singers versus their untrained counterparts. The results of this study indicate that alternative normative expectations for the DSI may need to be taken into account when using the DSI with patients who have participated in directed vocal training, such as choral participation and voice/singing lessons.

Key Words: Dysphonia Severity Index–Singers–Maximum phonation time–Jitter–Phonational frequency range.

INTRODUCTION

The perceptual evaluation of voice is considered to be an essential aspect of the conventional voice diagnostic that has relevance to most voice-disordered patients and provides a global measure of vocal performance readily available to all clinicians.[1] Although perceptual evaluation of voice has obvious importance, there are several limitations associated with this method of assessment that clearly influence its clinical utility. These limitations include problems with scale validity and reliability, particularly for midscale (ie, mild to moderate) pathological voices; lack of credibility for medical-legal purposes; poorly defined and/or shifting definitions of severity; and the intrusive effects of voice and speech characteristics other than the quality dimension that is meant to be judged.[1–3] Many of these limitations stem from the attempt to describe the voice via a temporary auditory impression of the acoustic signal. As a response, voice clinicians and researchers have added to the perceptual assessment of voice quality with other methods that provide a permanent record of the vocal behavior and allow for a more objective analysis of the patient's voice quality. Acoustic methods of voice analysis have been primary tools of both the clinician and the researcher for many years. These methods have become widely used in both research and clinical situations since the advent of relatively low-cost personal computers and analog-to-digital acquisition hardware in the early 1990s, and have the benefits of being noninvasive; readily available at relatively low cost compared with other methods of voice analysis; applicable to treatment and diagnosis; and are supported by a substantial body of literature.[4] Multiparameter acoustic models, which may be used to characterize voice function and quantify the severity of dysphonia, have been presented by Michaelis et al,[5] Callan et al,[6] Fröhlich et al,[7] and Awan and Roy.[8–10]

An alternative method for encompassing the multidimensional nature of the normal versus disordered voice is the Dysphonia Severity Index (DSI). The DSI was developed by Wuyts et al,[11] with the purpose of developing an index that would both objectively and quantitatively correlate with perceived voice quality. The DSI makes use of a combination of several voice measures that may be obtained from voice-assessment procedures, such as the voice range profile and basic aerodynamic and acoustic analyses: the highest phonational frequency (F_0 high in Hz), lowest intensity (I low in dB), maximum phonation time (MPT in seconds), and jitter (%). Because voice has been described as a multiparameter behavior, it appears reasonable that a multiparameter model, such as that presented in the DSI, may be useful in describing vocal function.[12] The components of the DSI form a specific combination of acoustic voice measures that may aid in characterizing various types of vocal dysfunction. Wuyts et al[11] stated that, when extra mass is evenly distributed along the true vocal fold(s), the higher vibratory rates become dampened. The result is a decrease in the upper reaches of the phonational frequency range. Structural changes to the true vocal folds, such as distributed or organized mass lesions, may increase glottal resistance such that greater subglottal pressures will be necessary to initiate and maintain vocal fold vibration. Consequently, the lowest phonational intensity will often be increased. In addition to changes in vocal intensity, vocal fold pathology (eg, unilateral or bilateral organized lesions or distributed tissue change; organic pathology affecting the ability to effectively tense or approximate the folds during phonation) often results in disturbances in the periodicity of phonation. These disturbances may be described in terms of jitter, a measure of short-term instability that quantifies cycle-to-cycle variations in frequency and has been used to assess the degree of perturbation in the voice signal. Finally, MPT has

Accepted for publication April 2, 2009.
Portions of this article were presented at the 37th Annual Symposium: Care of the Professional Voice, Philadelphia, PA, in May 2008.
From the *Department of Audiology & Speech Pathology, Bloomsburg University of Pennsylvania, Bloomsburg, Pennsylvania; and the †Department of Rehabilitation Sciences, College of Health Sciences, University of Kentucky, Lexington, Kentucky.
Address correspondence and reprint requests to Shaheen N. Awan, Department of Audiology & Speech Pathology, Bloomsburg University of PA, Centennial Hall, 400, East Second Street, Bloomsburg, PA 17815-1301. E-mail: sawan@bloomu.edu
Journal of Voice, Vol. 24, No. 6, pp. 661-666
0892-1997/$36.00

doi:10.1016/j.jvoice.2009.04.001

been regarded as a general measure of phonatory ability[13] that reflects the function of several mechanisms necessary for voice production, such as respiratory capacity and control, subglottic pressure, airflow resistance, and closure of the vocal folds. The DSI was initially described and validated in a study by Wuyts et al,[11] in which the authors obtained various acoustic and commonly used aerodynamic measures from 387 subjects (68 normal controls vs 319 voice-disordered subjects). In addition, each patient's voice was perceptually rated using the grade, roughness, breathiness, asthenia, strain (GRBAS) scale.[13] The DSI was obtained via multiple regression analyses and consists of four weighted variables in the equation: DSI $= 0.13 \times$ MPT (seconds) $+ 0.0053 \times F_0$ high (Hz) $- 0.26 \times I$ low (dB) $- 1.18 \times$ jitter (%) $+ 12.4$. Results of the study indicated an inverse relationship between the DSI and the grade (overall severity) of dysphonia, as well as between the DSI and the Voice Handicap Index (VHI). The DSI was transformed such that a DSI = +5 corresponded to G_0 (normal voice), and a DSI = -5 corresponded to G_3 (severe dysphonia). It was noted by these authors that the DSI is not necessarily restricted to the +5 to -5 range. A DSI of +1.6 was determined to be the cutoff for perceptually normal voices.

Several studies have used the DSI to objectively describe normal versus disordered voice characteristics and change in voice over time. Timmermans et al[14] evaluated voice quality change in 68 students (49 received voice training; 19 served as an untrained control group). The vocally trained group was provided with instruction regarding relaxation, posture, breathing pattern, and active articulation. Results showed a significant increase in the DSI for the trained group (from 2.3 to 4.5) versus no significant change in the untrained group. In a second study, Timmermans et al[15] used the DSI as a method of evaluating the effectiveness of a voice-training program for 23 professional voice users. The DSI scores were again observed to significantly increase from the time of training onset (mean DSI = 2.0) to 9 months post-training onset (mean DSI = 3.7) and to 18 months post-onset (mean DSI = 4.6), with the most prominent voice-characteristic change being increases in F_0 high. Hakkesteegt et al[16] investigated the possible influence of age and gender on the DSI (69 females; 49 males between 20 and 79 years of age). Although significant differences between males and females were observed for F_0 high and MPT, the mean DSI between the genders was not significantly different (mean DSI: females = 4.3; males = 3.8). The DSI was observed to decrease significantly with age in both genders, primarily because of reductions in F_0 high and low intensity. These authors inferred that the DSI of a particular voice patient should be compared with appropriate normative data from comparable age and gender subjects. Woisard et al[17] examined the possible correlation between a French version of the VHI and quantitative methods of describing voice, including the DSI. In contrast to Wuyts et al,[11] no significant correlation between the VHI and the DSI was observed. Woisard et al[17] indicated that the DSI should be seen as a source of clinical information independent of the VHI. Hakkesteegt et al[18] reported on the interobserver and test-retest variability of the DSI. Thirty normal subjects were measured by two speech pathologists on three different days (approximate interval of 1 week between measurements). The interobserver variability in measurement was observed to be low with very little influence on DSI variability. In addition, any differences in DSI measurement between the observers were reported as nonsignificant. These authors determined that intrasubject DSI changes needed to exceed 2.49 to be significant. Hakkesteegt et al[19] investigated the possible relationship between the GRBAS scale and the DSI. The subjects included 294 voice-disordered patients and 118 normal controls. Significantly lower DSI and higher grade (overall severity) scores were observed in disordered versus control groups. In addition, a DSI score of 3.0 was observed to discriminate between control and disordered groups with sensitivity = 0.72 and specificity = 0.75. It was observed that DSI scores may be reduced even for patients with overall grade/severity scores = 0, indicating that the voice complaints of some patients may not be solely quality based. Most recently, Hakkesteegt et al[20] examined the possible relationship between the DSI and the VHI. Pre- and postintervention measures were obtained from 171 voice-disordered patients. The subjects were divided into voice therapy, surgical intervention, and no intervention groups. Consistent with Woisard et al,[17] results indicated that the DSI and VHI measure different aspects of a voice disorder, with the VHI being a measure of patient perception and the DSI a measure of vocal performance/capability. Although both methods were able to show differences between pre- and postintervention groups, these authors indicated that DSI and VHI are not necessarily related.

The DSI has been reported to be a valuable clinical tool for the quantitative description of normal versus disordered voice. However, as indicated by Hakkesteegt et al,[16] extended normative data that will provide focused comparisons for subgroups of the normal population are necessary for appropriate clinical interpretations. A subgroup of the normal population that may be seen for voice complaints is that of the trained singer. Voice training has been said to influence the morphology and the control over the voice source.[21] Although inconclusive, several research studies have reported that trained singers may have increased respiratory capacities and different respiratory postures as compared with untrained subjects;[22–25] may have greater F_0 ranges than the normal untrained population;[26–28] and greater dynamic range capability.[27,29,30] In addition, Sulter et al[29] speculated that trained singers may have developed improved breath control during voicing, resulting in the ability to produce vocal fold oscillation at lower subglottal pressures. Whether these reported differences between trained and untrained singers are specifically the result of training itself; the regular participation in experiences such as choral singing; or to inherent physiological differences in those who choose to regularly participate in directed singing, is unclear. However, with these possible differences in mind, it would appear that DSI values for trained singers may be substantially different from those reported for untrained participants. Because of possible increased vocal capabilities, it may be that a trained singer who is experiencing some aspect of vocal dysfunction could still obtain a DSI score within the expected values reported by Wuyts et al.[11] The availability of data from trained singers