Acoustic Perturbation Measures Improve with Increasing Vocal Intensity in Individuals With and Without Voice Disorders

*M. Brockmann-Bauser, *J.E. Bohlender, and †,‡,§D.D. Mehta, *Zurich, Switzerland, and †‡§Boston, Massachusetts

Summary: Objective. In vocally healthy children and adults, speaking voice loudness differences can significantly confound acoustic perturbation measurements. This study examines the effects of voice sound pressure level (SPL) on jitter, shimmer, and harmonics-to-noise ratio (HNR) in adults with voice disorders and a control group with normal vocal status.

Study Design. This is a matched case-control study.

Methods. We assessed 58 adult female voice patients matched according to approximate age and occupation with 58 vocally healthy women. Diagnoses included vocal fold nodules (n = 39, 67.2%), polyps (n = 5, 8.6%), and muscle tension dysphonia (n = 14, 24.1%). All participants sustained the vowel /a/ at soft, comfortable, and loud phonation levels. Acoustic voice SPL, jitter, shimmer, and HNR were computed using *Praat*. The effects of loudness condition, voice SPL, pathology, differential diagnosis, age, and professional voice use level on acoustic perturbation measures were assessed using linear mixed models and Wilcoxon signed rank tests.

Results. In both patient and normative control groups, increasing voice SPL correlated significantly ($P < 0.001$) with decreased jitter and shimmer, and increased HNR. Voice pathology and differential diagnosis were not linked to systematically higher jitter and shimmer. HNR levels, however, were statistically higher in the patient group than in the control group at comfortable phonation levels. Professional voice use level had a significant effect ($P < 0.05$) on jitter, shimmer, and HNR.

Conclusions. The clinical value of acoustic jitter, shimmer, and HNR may be limited if speaking voice SPL and professional voice use level effects are not controlled for. Future studies are warranted to investigate whether perturbation measures are useful clinical outcome metrics when controlling for these effects.

Key Words: Acoustic perturbation–Harmonics-to-noise ratio–Voice diagnostics–Voice loudness–Occupational voice use.

INTRODUCTION

Instrumental measurements of acoustic perturbation form part of a comprehensive voice examination and are used to objectively describe vocal output.[1–3] The clinical application is based on the assumption that pathological changes in vocal fold mass or tension lead to increased and measurable irregularity or noise in the human voice signal.[4] For example, techniques such as videolaryngostroboscopy often restrict typical tongue movement during voice assessment. In addition, auditory-perceptual evaluations of voice are based on subjective ratings of vocal quality that are prone to psychometric reliability issues. In turn, instrumental indices, such as perturbation measurements, provide objective information about vocal output during natural voice and speech production using computer-assisted analyses of the acoustic speech signal.[1]

The present work focuses on the following widely applied acoustic perturbation measures: jitter, shimmer, and harmonics-to-noise ratio (HNR).[4] Jitter and shimmer are typically computed in the time domain and indicate variations in the cycle-to-cycle period duration and amplitude, respectively, across acoustic cycles during voice production. HNR can be computed in the time and spectral domains and indicates a ratio of harmonic energy to noise energy in the acoustic speech signal.[5] Despite a wide application to characterize voices with pathologies and to evaluate intervention success, the reliability and validity of acoustic perturbation measures are limited to date.[4,6,7] This has led to an uneven application of acoustic perturbation measures in clinical studies. Whereas organizations such as the American Speech-Language-Hearing Association are recommending supplanting jitter and shimmer measures with more robust acoustic metrics such as cepstral peak prominence,[8] some clinical research groups are using and further developing acoustic indices incorporating jitter and shimmer measures.[9–12]

Comparisons between groups of older adults and younger adults have shown age-related effects on vocal perturbation.[13,14] Also, in a meta-analysis of five studies with a total number of 51 adults between 21 and 80 years of age, jitter and shimmer tended to gradually increase with age.[15] However, in a study of 48 men between 25 and 75 years of age, jitter and shimmer were lowest in subjects in good physical condition, irrespective of age.[16] This result is supported by a recent study that demonstrated in 72 vocally normal adults that frequent voice training by singing attenuated aging effects on most acoustic parameters including fundamental frequency (f_o), mean voice sound pressure level (SPL), jitter, shimmer, and HNR.[17]

Accepted for publication April 17, 2017.

From the *Department of Phoniatrics and Speech Pathology, Clinic for Otorhinolaryngology, Head and Neck Surgery, University Hospital Zurich, University of Zurich, Frauenklinikstrasse 24, 8091 Zurich, Switzerland; †Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, One Bowdoin Square, 11th Floor, Boston, Massachusetts; ‡Department of Surgery, Harvard Medical School, Boston, Massachusetts; and the §MGH Institute of Health Professions, Boston, Massachusetts.

Address correspondence and reprint requests to M. Brockmann-Bauser, Department of Phoniatrics and Speech Pathology, Clinic for Otorhinolaryngology, Head and Neck Surgery, University Hospital Zurich, University of Zurich, Frauenklinikstrasse 24, 8091 Zurich, Switzerland. E-mail: meike.brockmann-bauser@usz.ch

Journal of Voice, Vol. 32, No. 2, pp. 162–168
0892-1997

https://doi.org/10.1016/j.jvoice.2017.04.008

Also, training effects on acoustic measurements of f_o, jitter, shimmer, and HNR have been shown in specific profession types such as high professional voice users or elite vocal performers.[18–20] To date, it is unclear whether effects of voice training are translated to habitual speaking voice characteristics in trained singers.[21] There is a possibility that underlying training effects have not been comprehensively described and therefore may influence the clinical measurement of acoustic voice perturbation.

In clinical measurements, usually patients are asked to produce sustained phonation of the vowel /a/, /i/, or /u/ with "comfortable pitch and loudness."[4,7,22] Under these measurement conditions, vowel effects have been documented in a number of works in individuals with and without voice disorders. For this reason, the current recommendation is to use the standard vowel /a/ in clinical practice.[7,22–25]

Whereas vowel effects may be relatively easy to control for in clinical assessments, the large natural differences in habitual speaking pitch and loudness present a more complex pragmatic problem.[26,27] Differences in speaking voice pitch (f_o) and loudness (voice SPL) have been shown to significantly affect measurements of jitter and shimmer in vocally healthy individuals.[4,22,28] Usually, we expect a natural covariation of voice f_o and SPL in measurements of speaking voice range profiles, with an association of higher voice SPL and increased f_o.[29–31] Videolaryngoscopic and aerodynamic examinations in healthy adults show that this is related with an increased vocal fold tonus.[32,33] A higher tonus might result in vocal fold stiffening, facilitating more regular vibration patterns and probably lower jitter and shimmer.[34] Thus, also jitter and shimmer and probably other indices of perturbation may show a natural covariation with voice SPL. This has been demonstrated by Pabon mapping acoustic perturbation results into voice range profile measurements, and might also apply to individuals with vocal pathology.

In a study of the proportional effects of vowel, gender, f_o, and voice SPL on jitter and shimmer in 57 vocally healthy adults, voice SPL was the largest influencing factor and accounted for up to 62% of the variation in shimmer. The effects of gender, vowel, and f_o accounted for up to 6% of measurement differences and thus were statistically smaller by comparison.[22] To date, it is not clear if these effects also apply to other indices of vocal perturbation or irregularity such as HNR. Also, this relation has been investigated only in vocally healthy adults and children.[4,7,22,23,35] Therefore, the main aims of the present work were to study SPL-related effects on jitter, shimmer, and HNR in individuals with and without diagnosed voice disorders, while also considering the influence of age and occupation-related voice use level.

METHODS

Subject sample and inclusion criteria

In a retrospective matched case-control study, 116 adult women aged between 18 and 64 years were drawn from a larger project studying ambulatory voice monitoring.[36] The present study extracted laboratory voice recordings from 58 adult female patients diagnosed with phonotraumatic vocal hyperfunction (vocal fold nodules or polyps) or non-phonotraumatic vocal hyperfunction (muscle tension dysphonia [MTD]) before and, in some cases, after treatment. Diagnoses included vocal fold nodules (n = 39, 67.2%), polyps (n = 5, 8.6%), and MTD (n = 14, 24.1%). Each patient was paired with a vocally healthy control subject who was matched according to sex, approximate age (±5 years), and occupation (profession).

Diagnoses were based on a complete team evaluation by laryngologists and speech-language pathologists at the Massachusetts General Hospital Voice Center including (1) a case history, (2) endoscopic imaging of the larynx, (3) aerodynamic and acoustic assessment of vocal function, (4) patient-reported Voice-Related Quality of Life questionnaire, and (5) clinician-administered Consensus Auditory-Perceptual Evaluation of Voice assessment. Normal voice status of the vocally healthy participants was confirmed via interview and a laryngeal stroboscopic examination. Of the included 58 patients, 33 patients had voice assessments before and after laryngeal surgery or voice therapy. Informed consent was obtained from all subjects, and all experimental protocols were approved by the institutional review board of Partners HealthCare System at Massachusetts General Hospital.

Subjects with voice disorders had a mean age of 27.8 years (18–64 years, standard deviation [SD]: 12.1 years), and the matched-control subjects with normal voices had a mean age of 27.8 years (18–61 years, SD: 11.8 years). As determined by a linear mixed model (LMM) analysis, there was no statistical difference in age distribution between the two groups ($P > 0.05$).

Table 1 displays a classification of each profession into four subgroups according to voice use level after Koufman and Isaacson,[37] modified by do Amaral Catani et al who reclassified teachers as level II (*versus* level III) voice users.[38] For the current study, 35 subject pairs were elite vocal performers (level I voice use level), 10 pairs were professional voice users (level II), 8 pairs were non-vocal professionals (level III), and five pairs were non-vocal non-professionals (level IV) (Table 1).

Acoustic recording technique and protocol

Acoustic voice recordings were acquired using a head-mounted microphone integrated in a pneumotachograph mask in an off-axis position at a distance of 10 cm from the lips (MKE 104, Sennheiser, Electronic GmbH, Wennebostel, Germany). The microphone signal was input to a preamplifier (model 302 Dual Microphone Preamplifier, Symetrix, Inc., Mountlake Terrace, WA), followed by preconditioning electronics (CyberAmp 380, Axon Instruments, Inc., Union City, CA) for gain control and anti-alias filtering at a 3-dB cutoff frequency of 8 kHz. The analog signal was digitized at a 20-kHz sampling rate, 16-bit quantization, and ±10-V voltage range (Digidata 1440A, Axon Instruments, Inc.). All subjects were asked to sustain a prolonged vowel /a/ at a comfortable pitch in their typical speaking voice mode at an individually "soft," "comfortable," and "loud" voice intensity level.

Analysis technique and main outcome measures

Each acoustic signal was perceptually examined for instability and visually displayed using *Praat* (version 5.4.1.4; http://www.praat.org/) with an oscillogram and "Show intensity" and

TABLE 1.
Classification of Professions for Each Subject Pair According to Voice Use Level After Koufman and Isaacson, Modified by do Amaral Catani et al[37,38]

Level	Description	Number of Subject Pairs
Level I	Elite vocal performer, for whom even slight vocal difficulty may cause serious consequences. In the present study, these included professional singers, actors, and voice students.	35
Level II	Professional voice user, for whom moderate vocal difficulty would prevent adequate job performance. These included teachers and sports instructors.	10
Level III	Non-vocal professional, for whom severe dysphonia would prevent adequate job performance. These include doctors, social workers, psychologists, and business persons.	8
Level IV	Non-vocal non-professional, for whom suffering from vocal difficulties would not prevent adequate job performance. These include administrators, librarians, and college or university students.	5

"Show pulses" settings turned on.[5] Excluded were all recordings with Type 2 and Type 3 signals, incorrect or unstable f_o and voice SPL, signal clipping, or phonation time <1.5 seconds.[39] These criteria led to the inclusion of unequal numbers of voice recordings per loudness level and in patients before and after treatment (Table 3). Each recording was edited into an individual sound file using *Praat*. To exclude the increased variability of the voice onset and offset phase, only the signal segment from 0.5 seconds to 1.0 seconds from voice onset was acoustically analyzed. Calibrated voice SPL levels were obtained using the comparison method with a complex tone stimulus of known SPL.[40]

Table 2 lists the main outcome measures from the instrumental acoustic analysis performed with a custom *Praat* analysis script: voice SPL (dB SPL), jitter (%), shimmer (%), HNR (dB), and f_o (Hz). Jitter and shimmer were chosen, because both were normalized for an individual's voice SPL and f_o. As discussed in the introduction, a natural covariation of f_o with voice SPL was expected; therefore, f_o was also measured. Because this variable was not manipulated by task choice, only descriptive data were included (Table 3).

Statistical analysis

Data were coded in Excel and analyzed with *SPSS Statistics* version 22 (IBM Corp., Armonk, NY). First, descriptive statistics of the mean, SD, minimum, maximum, and range were computed for the acoustic outcome measures voice SPL, jitter, shimmer, and HNR. Because repeated measurements tend to be more similar within individuals than across individuals, linear mixed models (LMM) were used to investigate the overall effects of categorical voice intensity level (soft, comfortable, loud), continuous voice SPL (dB SPL), presence of pathology (absence or presence), differential diagnosis (nodules, polyps, MTD), professional voice use level (Levels I–IV), and age (continuous variable) on jitter, shimmer, and HNR.[41]

Further, because the study sample consisted of naturally matched pairs, the nonparametric paired Wilcoxon test was used to test statistical differences between the acoustic outcome measures from the patient and the control groups, and within the patient group before and after treatment. Jitter and shimmer were transformed logarithmically to statistically stabilize their large naturally observed measurement variance. Results of the statistical analysis were considered significant at $P < 0.05$.

RESULTS

Acoustic outcome measures per phonation level

Table 3 reports descriptive statistics for each acoustic measure for the two subject groups, including pre- and post-treatment assessments for the patient group. There was a statistically significant difference in voice SPL for the soft, comfortable, and loud conditions within the patient and normal subject groups ($P < 0.001$). As expected, mean f_o increased with voice SPL and was significantly different for each of the three loudness conditions ($P < 0.05$).

Mean voice SPL in comfortable phonation was 87.7 dB SPL (SD: 5.6 dB, range: 71.0–96.7 dB) for the normative group and,

TABLE 2.
Analyzed Instrumental Acoustic Parameters With Abbreviations, Units, Description, and Labels as Applied in *Praat* Software

Outcome Measure	Unit	Description and Comments	*Praat* Software Label
Voice SPL	dB SPL	Calibrated voice SPL values were determined using the comparison method	Mean energy intensity
Jitter	%	Relative cycle-to-cycle deviation from mean cycle period	Jitter (local)
Shimmer	%	Relative cycle-to-cycle deviation from mean cycle amplitude	Shimmer (local)
HNR	dB	Degree of periodicity in an acoustic signal, where HNR of 0 dB indicates an equal energy distribution in harmonic and noise signal components	Mean harmonics-to-noise ratio
f_o	Hz	Fundamental frequency	Median pitch

TABLE 3.
Descriptive Statistics for Voice SPL, Jitter, Shimmer, HNR, and f_o Within Each Loudness Condition for the Subject Groups With and Without Voice Disorder

Acoustic Parameter		Control Group With Normal Voices			Patient Group Before Treatment			Patient Group After Treatment		
		Soft	Comfortable	Loud	Soft	Comfortable	Loud	Soft	Comfortable	Loud
Calibrated SPL (dB SPL)	Mean	81.1	87.7	95.8	79.5	88.0	95.9	77.9	89.0	97.3
	SD	6.0	5.6	4.7	5.4	4.5	4.3	5.2	4.7	4.6
	Minimum	66.1	71.0	85.5	68.2	77.6	86.4	63.3	79.3	88.0
	Maximum	95.5	96.7	105.8	93.2	99.5	106.7	87.2	98.5	107.3
	Range	29.4	25.7	20.3	25.0	21.9	20.4	23.8	19.2	19.3
Jitter (%)	Mean	0.38	0.30	0.24	0.41	0.32	0.24	0.38	0.29	0.21
	SD	0.20	0.19	0.13	0.21	0.13	0.12	0.17	0.16	0.09
	Minimum	0.11	0.10	0.09	0.18	0.16	0.10	0.08	0.14	0.10
	Maximum	1.34	1.36	0.91	1.06	0.72	0.85	0.96	0.79	0.59
	Range	1.22	1.26	0.81	0.88	0.56	0.76	0.88	0.65	0.48
Shimmer (%)	Mean	2.66	1.65	1.19	2.74	1.97	1.32	2.70	1.67	1.05
	SD	1.31	0.74	0.64	1.32	1.20	0.66	1.52	0.81	0.46
	Minimum	1.14	0.65	0.28	1.17	0.71	0.43	0.68	0.70	0.46
	Maximum	7.17	3.94	3.00	9.23	7.84	4.58	9.49	5.17	2.93
	Range	6.03	3.29	2.73	8.07	7.13	4.15	8.81	4.46	2.47
Mean HNR (dB)	Mean	25.1	27.7	29.8	24.4	26.5	29.4	24.9	28.6	30.6
	SD	3.7	3.2	2.9	4.1	3.4	3.3	4.8	2.4	2.1
	Minimum	16.5	20.8	23.8	11.0	14.9	21.9	15.4	23.6	25.2
	Maximum	31.8	33.3	36.0	34.8	34.0	35.1	39.6	33.5	34.7
	Range	15.4	12.5	12.2	23.8	19.1	13.2	24.2	9.9	9.5
Mean f_o (Hz)	Mean	244.1	249.2	266.6	248.4	243.3	253.4	255.3	253.0	265.5
	SD	41.2	36.5	43.6	43.9	41.9	37.8	30.9	28.3	33.8
	Minimum	162.0	172.0	179.6	138.3	154.7	189.2	205.7	202.9	206.6
	Maximum	317.6	318.3	368.8	379.2	381.1	379.9	317.2	306.7	351.2
	Range	155.6	146.3	189.1	240.9	226.4	190.8	111.6	103.8	144.6
Total n		52	58	54	53	57	57	32	31	33

The number of included recordings is indicated by “total n” and varied per loudness condition within each subject group.

similarly, 88.0 dB SPL (SD: 4.5 dB, range: 77.6–99.5 dB) for the patient group (Table 3). There was no significant difference in mean voice SPL between the patient and the control groups within the three phonation levels (soft, comfortable, and loud) according to LMM and Wilcoxon signed rank analyses ($P > 0.641$).

Effect of loudness condition and voice SPL

Both categorical loudness condition (soft, comfortable, loud) and calibrated voice SPL (dB SPL) had a highly significant effect on jitter, shimmer, and HNR across the normative and patient groups ($P < 0.001$). Figure 1 shows within-group univariate relationships between the acoustic perturbation measures and voice SPL. Jitter decreased, shimmer decreased, and HNR increased, indicating an overall improvement in these perturbation measures with increasing voice SPL. The regression line indicates a potential mathematical correction for comparing jitter, shimmer, and HNR across different voice SPL values. This type of correction by applying R^2 is tempered, however, by the large natural data variability around the regression lines within both normative and patient voice samples.

Effect of presence and type of pathology

Using LMM analysis, we found no statistically significant differences in jitter, shimmer, and HNR between the patient and the normative control groups or among diagnoses in the patient group ($P = 0.097$–0.525) with respect to loudness condition and voice SPL. Also, there was no interaction between the presence of pathology and voice SPL for all investigated instrumental parameters ($P = 0.053$–0.771). Even though only suitable voice signals were chosen for analysis, Figure 1 indicates that these results may have been influenced by several outliers in the control group.

However, using the nonparametric Wilcoxon test that takes advantage of the patient-control pairings, there was a statistically significant difference for HNR between the patient and the normative groups in the comfortable loudness condition ($P = 0.01$). Also, there were significant differences in HNR between patient measures before and after treatment in the comfortable loudness condition ($P = 0.004$). The HNR in the control group (mean: 27.7 dB, SD: 3.2 dB) was 1.2-dB greater relative to the measurements in the pretreatment patient group (mean: 26.5 dB, SD: 3.4 dB). Furthermore, the large observed spread of 12.5 dB in the control and 19.1 dB in the patient group shows

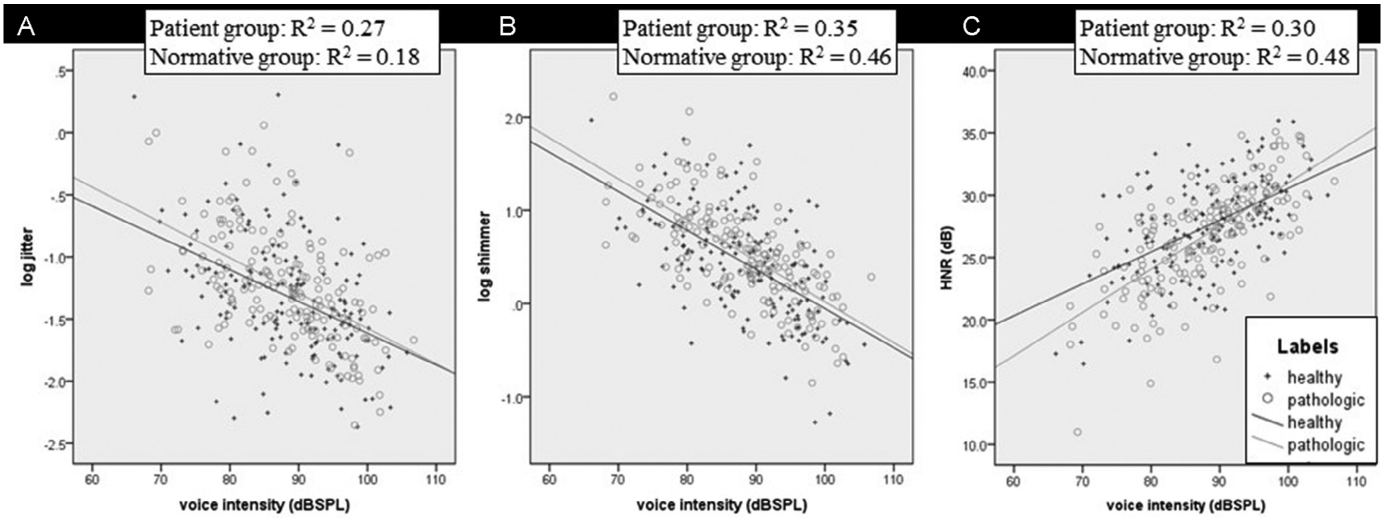

FIGURE 1. Voice SPL effects on jitter, shimmer, and HNR in healthy and pathological voices. Improvement of (A) jitter (*left*), (B) shimmer (*middle*), and (C) HNR (*right*) with increasing voice SPL within the normative (*black crosses*) and pretreatment patient (*gray circles*) data. The bright gray line indicates the regression line for the normative group, and the dark gray line indicates the regression line for the patient group. R^2 expresses the correction factor, which may be used to mathematically adapt jitter, shimmer, and HNR for a defined SPL level.

the limited clinical applicability of these results in voice diagnostics. There were no differences for HNR in the soft and loud conditions, pointing to the potential importance of controlling for voice intensity during clinical voice assessment. Jitter and shimmer measures were not statistically different between individuals with and individuals without a voice disorder.

Effect of age and professional voice use level

As determined by LMM analysis, age did not have a statistical effect on the calculation of jitter, shimmer, and HNR. However, these acoustic perturbation measures were significantly different depending on an individual's professional voice use level; that is, when controlling for loudness condition (soft, comfortable, loud), statistically significant differences were exhibited among the four professional voice use levels for jitter ($P = 0.005$), shimmer ($P = 0.017$), and HNR ($P < 0.001$). Notably, professional voice use level affected voice SPL only in the loud condition in both normative ($P = 0.042$) and patient ($P = 0.08$) groups, but not in the soft and comfortable loudness conditions ($P > 0.05$). When considering voice SPL as continuous variable, we found only a significant professional voice use level effect for jitter ($P = 0.046$) and HNR ($P = 0.003$).

DISCUSSION

In clinical measurements of jitter, shimmer, and HNR, the patient's individual speaking voice SPL is a significant confounding factor. Regardless of the presence of a voice disorder, there was an improvement in jitter, shimmer, and HNR with increasing voice loudness. The observed confounding voice SPL effects may also affect other variants of acoustic perturbation (calculated by algorithms other than those applied in the present work) and acoustic analysis strategies, such as the Goettingen Hoarseness Diagram,[42] Dysphonia Severity Index,[43] and Acoustic Voice Quality Index.[9]

Are age and profession relevant factors in clinical measurements?

In our sample of adult women between 18 and 64 years, age did not affect jitter, shimmer, and HNR measurements. However, jitter and shimmer may also reflect the general physical condition, irrespective of age.[16] In the present work, 60% of participants were level I elite vocal performers who often are vocally trained.[37] In singers, changes in a number of acoustic parameters—including jitter, shimmer, f_o, and voice SPL—were explained with a better production and control of vocal fold tonus.[18,19,44,45] There were significant differences between profession groups in subjectively loud phonations, which support this hypothesis. This observation highlights that occupation-related effects may be partially caused by underlying voice SPL differences between groups for specific voice tasks. Differences in voice training experience should thus be considered as a relevant factor estimating acoustic voice measures. Future studies in a larger clinical sample that includes both women and men are warranted to aid in defining robust markers of voice training status.

Implications for phonation models

The acoustic perturbation measures improved with increasing voice SPL in both patient and normative subjects. This result may indicate an underlying physiological mechanism, perhaps analogous to the known covariation between voice SPL and f_o, which was also observed in the present study.[29] Videolaryngoscopic and aerodynamic examinations in adults with normal voices have demonstrated that increases in voice SPL and f_o are associated with higher vocal fold tonus.[32,33] This might lead to a stiffer and stabilized vocal fold, thereby reducing random variability in vocal fold vibratory patterns.[34] Further, as discussed above, it has been proposed that training effects may lead to better production and control of vocal fold tonus.[18,19,44,45] In that view, the significant influence of professional voice use level—hence, formal *versus* nonformal voice training—supports this proposed hypothesis.

Consequences for the clinical application of perturbation parameters

In the present analysis of 58 women with voice disorder and 58 who are vocally healthy, matched after age and profession, there was no significant effect of pathology or diagnosis type on jitter and shimmer. These results may be partially explained by the choice of Type 1 recordings only and the elimination of training effects related to professional voice use through the study design.

For HNR, there was a significant difference between disordered and healthy voices at comfortable phonation levels, but only in comparisons between matched pairs. However, the clinical usefulness of these results is tempered by the comparatively small difference of 1.2 dB between healthy and pathological voice samples and the observed large overall data spread. Thus, HNR may provide better clinically relevant information than jitter and shimmer, but only under sufficient control of profession and voice SPL effects.

This leads us to a key question in clinical practice: How do we control for the observed significant voice SPL effects practically? Whereas vowel effects are comparatively easy to control (by simply asking all patients to use the same vowel), the answer for voice SPL effects is more complex. As discussed by Brown and colleagues,[26,27] speakers respond with different voice intensities when asked to phonate with "comfortable" voice loudness in identical environments. In our sample, voice intensity for the comfortable loudness condition spanned 71.0–96.7 dB SPL in the normative subject group and 77.6–99.5 dB SPL in the patient group.

One way to control for these effects is to modify the clinical voice task. As shown in a previous work, vocally healthy children (>6 years of age) and adults were able to control their own voice SPL by using visual feedback.[23] In a study of 20 vocally healthy women and 20 vocally healthy men, subjects were asked to phonate at 65, 75, 85, and 95 dBA (recording distance 10 cm) and provided with visual feedback. The most accurate SPL was produced for the task to phonate at 85 dBA. Under these conditions, women produced a mean of 85.3 dBA (SD: 2.7 dB) and men a mean of 84.8 dBA (SD: 1.3 dB).[46] However, in a clinical examination situation, there are clear pragmatic, ethical, and legal considerations to weigh. First, some patients may not be able to produce such a voice intensity level. And, even if patients were able to perform this task, in some organic disorders such as acute vocal fold inflammation, it may be painful or even harmful to phonate at around 85 dBA. Furthermore, phonating at a prescribed level might not best reflect habitual vocal behavior, which is presumed to contribute to the voice disorder.

Another way to control for SPL-related effects may be by using a correction factor or formula. As implied by the linear regression results of Figure 1, it could be possible to apply this statistical correction for jitter, shimmer, and HNR as a function of voice SPL. In this way, all parameters could be statistically adapted to a specified voice intensity level. However, it is recognized that there may be an inherently large measurement variance in the data across voice SPL in both patient and matched-control groups, hindering the discrimination of healthy from pathological acoustic signals. As discussed in detail by Ternström and colleagues, voice function and hence the produced acoustic voice signal is influenced by numerous internal and external sources of variability.[31] For example, voice perturbation has also been shown to vary with the vocal register, even in phonations with stable f_o.[30,47,48] In summary, all proposed approaches would require further studies in confounding factors, how they behave as a function of voice SPL and f_o, and how to best control for them to obtain clinically useful perturbation measurements.

This leads to the general question of which clinical information we specifically search by using acoustic parameters. As discussed in the introduction, measurements of vocal irregularity may provide information about voice function that is not available from other assessment techniques. If the aim were to assess typical voice behavior, the strategy to control for voice SPL effects using prescribed levels would not be optimal. Another option is to work with a standardized set of a variety of voice tasks, such as in vocal loading tests, and to map the resulting acoustic perturbation measurements into voice range profiles.[30,49] In turn, if the aim were to detect discrete organic lesions, we need more robust clinical evidence that controlling for voice intensity results in clinically significant differences between patient and normative groups.

CONCLUSIONS

An individual's vocal loudness level may act as a significant confounding factor during clinical voice assessment when estimating acoustic perturbation measures. Overall, acoustic perturbation improved as voice intensity increased, with jitter and shimmer decreasing and HNR increasing. Similar effects may also apply to other acoustic voice measures that use or combine further jitter, shimmer, or HNR parameter types. Furthermore, an individual's professional voice use level may influence these acoustic voice measures as well, pointing toward potential training effects. Future studies are warranted to investigate clinically useful acoustic outcome measures after adequately controlling for voice intensity and occupation.

Acknowledgments

The authors acknowledge Dr Jarrad H. Van Stan for subject recruitment and data collection; Melissa Cooke, Amanda Fryd, and Molly Bresnahan for help with signal segmentation; and Dr Robert E. Hillman for study support and interpretation.

The authors thank Dr Malzorgata Roos from the Department of Biostatistics, University of Zurich, Switzerland, for assistance with the statistical analysis of the data, and Dr Sandra Schwab and Dr Lei He from the Phonetics Laboratory of the University of Zurich, Switzerland, for their input on the acoustic analysis procedure.

This work was supported in part by the National Institutes of Health National Institute on Deafness and Other Communication Disorders under Grant R33 DC011588 (PI: Hillman) and in part by the Voice Health Institute. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

For this work, M. Brockmann-Bauser received the Hamdan International Presenter Award 2016 from The Voice Foundation, Philadelphia, PA, USA.

REFERENCES

1. Dejonckere P, Bradley P, Clemente P, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *Eur Arch Otorhinolarygol*. 2001;258:77–82.
2. Cohen W, Wardrop A, Wynne DM, et al. Development of a minimum protocol for assessment in the paediatric voice clinic. Part 2: subjective measurement of symptoms of voice disorder. *Logoped Phoniatr Vocol*. 2012;37:39–44.
3. Cohen W, Wynne DM, Kubba H, et al. Development of a minimum protocol for assessment in the paediatric voice clinic. Part 1: evaluating vocal function. *Logoped Phoniatr Vocol*. 2012;37:33–38.
4. Brockmann-Bauser M, Drinnan MJ. Routine acoustic voice analysis: time to think again? *Curr Opin Otolaryngol Head and Neck Surg*. 2011;19:165–170.
5. Boersma P, Weenink D. PRAAT. 5.4.14, Amsterdam: University of Amsterdam; 2015. Available at: http://www.fon.hum.uva.nl/praat/. Accessed August 4, 2015.
6. Carding PN, Wilson JA, Mackenzie K, et al. Measuring voice outcomes: state of the science review. *J Laryngol Otol*. 2009;123:823–829.
7. Brockmann-Bauser M, Beyer D, Bohlender JE. Reliable acoustic measurements in children between 5;0 and 9;11 years: gender, age, height and weight effects on fundamental frequency, jitter and shimmer in phonations without and with controlled voice SPL. *Int J Pediatr Otorhinolaryngol*. 2015;79:2035–2042.
8. Roy N, Barkmeier-Kraemer J, Eadie T, et al. Evidence-based clinical voice assessment: a systematic review. *Am J Speech Lang Pathol*. 2013;22:212–226.
9. Barsties B, Maryn Y. External validation of the acoustic voice quality index version 03.01 with extended representativity. *Ann Otol Rhinol Laryngol*. 2016;125:571–583.
10. Barsties VLB, Maryn Y, Gerrits E, et al. The Acoustic Breathiness Index (ABI): a multivariate acoustic model for breathiness. *J Voice*. 2017.
11. Hosokawa K, Barsties B, Iwahashi T, et al. Validation of the acoustic voice quality index in the Japanese language. *J Voice*. 2016.
12. Uloza V, Petrauskas T, Padervinskis E, et al. Validation of the acoustic voice quality index in the Lithuanian language. *J Voice*. 2016.
13. Schaeffer N, Knudsen M, Small A. Multidimensional voice data on participants with perceptually normal voices from ages 60 to 80: a preliminary acoustic reference for the elderly population. *J Voice*. 2015;29:631–637.
14. Stathopoulos E, Huber J, Sussmann J. Changes in acoustic characteristics of the voice across the life-span: measures from 4–93 year olds. *J Speech Lang Hear Res*. 2011;54:1011–1021.
15. Baken RJ, Orlikoff RF. *Clinical Measurements of Speech and Voice*. 2nd ed. Albany, New York: Thomson Delmar Learning; 2000.
16. Ramig LA, Ringel RL. Effects of physiological aging on selected acoustic characteristics of voice. *J Speech Hear Res*. 1983;26:22–30.
17. Lortie CL, Rivard J, Thibeault M, et al. The moderating effect of frequent singing on voice aging. *J Voice*. 2016.
18. Walzak P, McCabe P, Madill C, et al. Acoustic changes in student actors' voices after 12 months of training. *J Voice*. 2008;22:300–313.
19. Mendes AP, Brown WS Jr, Rothman HB, et al. Effects of singing training on the speaking voice of voice majors. *J Voice*. 2004;18:83–89.
20. Rothman HB, Brown WS Jr, Sapienza CM, et al. Acoustic analyses of trained singers perceptually identified from speaking samples. *J Voice*. 2001;15:25–35.
21. Brown WS Jr, Rothman HB, Sapienza CM. Perceptual and acoustic study of professionally trained versus untrained voices. *J Voice*. 2000;14:301–309.
22. Brockmann M, Drinnan MJ, Storck C, et al. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *J Voice*. 2011;25:44–53.
23. Brockmann-Bauser M, Beyer D, Bohlender JE. Clinical relevance of speaking voice intensity effects on acoustic jitter and shimmer in children between 5;0 and 9;11 years. *Int J Pediatr Otorhinolaryngol*. 2014;78:2121–2126.
24. Sussmann JE, Sapienza C. Articulatory, developmental, and gender effects on measures of fundamental frequency and jitter. *J Voice*. 1994;8:145–156.
25. Kiliç MA, Ögüt F, Dursun G, et al. Effects of vowels on voice perturbation measures. *J Voice*. 2004;18:318–324.
26. Brown WJ, Morris R, Murry T. Comfortable effort level revisited. *J Voice*. 1996;10:299–305.
27. Brown WS, Shrivatsav R. Comfortable effort level in young children's speech. *Folia Phoniatr Logop*. 2007;59:227–233.
28. Gelfer MP. Fundamental frequency, intensity, and vowel selection: effects on measures of phonatory stability. *J Speech Hear Res*. 1995;38:1189–1198.
29. D'Alatri L, Marchese MR. The speech range profile (SRP): an easy and useful tool to assess vocal limits. *Acta Otorhinolaryngol Ital*. 2014;34:253–258.
30. Pabon J. Objective acoustic voice-quality parameters in the computer phonetogram. *J Voice*. 1991;5:203–216.
31. Ternström S, Pabon P, Södersten M. The voice range profile: its function, applications, pitfalls and potential. *Acta Acustica*. 2016;102:268–283.
32. Hodge F, Colton R, Kelley R. Vocal intensity characteristics in normal and elderly speakers. *J Voice*. 2001;15:503–511.
33. Sulter AM, Schutte HK, Miller DG. Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. *J Voice*. 1996;10:175–189.
34. Laukkanen A, Ilomäki I, Leppänen K, et al. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
35. Brockmann M, Storck C, Carding PN, et al. Voice loudness and gender effects on jitter and shimmer in healthy adults. *J Speech Lang Hear Res*. 2008;51:1152–1160.
36. Mehta DD, Van Stan JH, Zanartu M, et al. Using ambulatory voice monitoring to investigate common voice disorders: research update. *Front Bioeng Biotechnol*. 2015;3:155.
37. Koufman JA, Isaacson G. The spectrum of vocal dysfunction. *Otolaryngol Clin North Am*. 1991;24:985–988.
38. do Amaral Catani GS, Hamerschmidt R, Moreira AT, et al. Subjective and objective analyses of voice improvement after phonosurgery in professional voice users. *Med Probl Perform Art*. 2016;31:18–24.
39. Titze IR. *Workshop on Acoustic Voice Analysis: Summary Statement*. Iowa City, IA: National Center for Voice and Speech; 1995.
40. Winholtz WT I. Conversion of a head-mounted microphone signal into calibrated SPL units. *J Voice*. 1997;11:417–421.
41. Cnaan A, Laird NM, Slasor P. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat Med*. 1997;16:2349–2380.
42. Frohlich M, Michaelis D, Strube HW, et al. Acoustic voice analysis by means of the hoarseness diagram. *J Speech Lang Hear Res*. 2000;43:706–720.
43. Wuyts FL, De Bodt MS, Molenberghs G, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res*. 2000;43:796–809.
44. Mendes AP, Rothman HB, Sapienza C, et al. Effects of vocal training on the acoustic parameters of the singing voice. *J Voice*. 2003;17:529–543.
45. Tay EY, Phyland DJ, Oates J. The effect of vocal function exercises on the voices of aging community choral singers. *J Voice*. 2012;26:672, e619-627.
46. Brockmann-Bauser M. *Improving Jitter and Shimmer Measurements in Normal Voices*. Idstein: Schulz-Kirchner Verlag; 2012.
47. Echternach M, Traser L, Richter B. Perturbation of voice signals in register transitions on sustained frequency in professional tenors. *J Voice*. 2012;26:674, e679-615.
48. Blomgren M, Chen Y, Ng ML, et al. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J Acoust Soc Am*. 1998;103:2649–2658.
49. Echternach M, Nusseck M, Dippold S, et al. Fundamental frequency, sound pressure level and vocal dose of a vocal loading test in comparison to a real teaching situation. *Eur Arch Otorhinolaryngol*. 2014;271:3263–3268.