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. ■■, No. ■■, pp. ■■-■■
0892-1997

http://dx.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