Abstract: This paper investigates vocal fold (VF) vibratory properties using quantitative analysis of high-speed digital imaging (HSDI) based on Nyquist-plot method derived from voicing during production of aberrant voice quality (VQ), clinically referred to as diplophonia, and defines the mechanism responsible for diplophonia and show how treatment (Tx) effects this VQ and VF behavior. In particular, pre- and post-Tx HSDI recordings of a female patient with muscular tension dysphonia (MTD) were analyzed using new quantitative analysis system for HSDI that involves tracing of VF edge and generation of glottal waveform and VF displacement, allowing us to define quantitative measures of vibratory symmetry and synchronization of VF vibrations, with subsequent analyses of glottal waveforms using Nyquist formula, to reveal vibratory pattern and characteristics of the vocal folds during this aberrant sound production, and later during normative phonation post Tx.

This is first ever HSDI and Nyquist-plot based analyses of aberrant voice known as diplophonia derived here from vocalization of a MTD case. The results reveal definitive and specific character of VF vibration responsible for this VQ.

Keywords: High-speed digital imaging, vocal-fold vibration, diplophonia, Nyquist plot

I. INTRODUCTION

Control of VF vibrations can be variably affected by organic and functional causation resulting in vibratory irregularity and deviant/aberrant acoustic product [1]. Such deviant vocal outputs have been analyzed in the past using various techniques including acoustics, aerodynamics and visualization, leading to improved understanding of how vocal fold physiology relates to the actual sound production and how to treat the underlying pathology. Usage of acoustics combined with visualization has been shown to provide improved information on the underlying pathology and underscores the value of multidisciplinary approach and clinical power [2], hence improving characterization of VF vibrations associated with some voice disorders, specifically for those that generate similar perceptual effects, confusing unequivocal clinical diagnosis that often is made by the ear alone.

To improve the analysis and to provide up-to-date explanation, newest visualization technique to study VF dynamics based on HSDI, also termed high speed videoendoscopy (HSV), combined with application of Nyquist-plot analysis [3], [4] were used here to study aberrant vocalization known clinically as diplophonia [1]. The HSDI system records images of the vocal folds at an acquisition rate of 2000 frames per second, fast enough to capture the vibration of the vocal folds in more details. To overcome the cumbersome subjective evaluation of analyzing such massive data, quantitative methods for HSDI-based and acoustic analyses have recently been developed [3-5], and were used here. These analyses generate comprehensive patterns of VF vibrations and are capable of defining the characteristics of the VF vibration in terms of robust, quantitative measures that enhance understanding of the mechanism of phonation not only in normative voice, but specifically in voice pathologies, which can not be handled easily by traditional stroboscopic illumination.

Phonatory VQ referred to in broad terms as “diplophonia, multiphonia or biphonation” is assumed to represent simultaneous production of at least two distinct tones during phonation. This pattern may be induced by an imbalance in bilateral tension or mass of the VF or by events within each VF, as evidenced in clinical cases representing mass, paralysis, neurological driven or functional dysphonias including non-true VF driven voice pathology. The mechanism underlying diplophonia has been studied by different researches using diverse experimental and theoretical approaches including biomechanical modeling, analysis of acoustic recordings and direct imaging of voice production. These studies suggest that diplophonia represents 1) an abnormal
glottic cycle with double or even multiple phased opening and closing; 2) asymmetric vibrations of the vocal folds, in which the left and right folds vibrate at a different frequency, or 3) out-of-phase vibratory pattern; 4) within cord variability and/or 5) combinations of the above. In this study we aim to provide a comprehensive, quantitative analysis of glottal area waveform (GAW) derived from HSDI recordings of diplophonic phonation from a clinical case representing idiopathic diplophonia, of MTD type.

II. METHODS

Data were acquired from HSDI recordings of a female patient with idiopathic diplophonia of MTD type, prior to and following voice therapy. In this way the subject was serving as her own control. Post data processing involved automated image segmentation, detection of VF edge and generation of GAW and VF displacements described by us previously [3-5]. Of specific interest here was the usage of Nyquist plot based analysis of the HSDI-derived waveforms and associated perturbation measures [3], [4]. This technology is used to generate characteristic patterns of the VF vibration and to describe quantitatively the VF vibration in the MTD patient’s voicing before and after treatment.

New measures of the VF vibration that include symmetry/ homogeneity, synchronization, and sustainability are defined. These measures along with the jitter metrics formed the basis of a robust quantitative analysis of the VF vibrations in pre- and post-Tx MTD voices. Further, the analysis provided an automatic, robust calculation of the glottic closure characteristics including the open quotient (OQ), speed quotient (SQ) and glottal closure index. Nyquist plot based waveform analyses and associated perturbation measures of the MTD voice were shown to generate not only at-a-glance patterns of the spatial- and temporal characteristic but also quantitative measures of the VF vibrations for the MTD voice.

III. RESULTS

HSDI-based quantitative analyses of the MTD voice demonstrated that the anterior and posterior of the VF undergo non-identical vibrations with distinct patterns of opening and closing during the diplophonic phase (Fig. 1). In particular, the anterior and medial portions of the VF exhibited more obvious bi-cyclic vibratory pattern compared to the posterior portion of the VF (data not shown). In contrast, vibrations within the left-right folds are almost symmetrical throughout the anterior-posterior (A-P) locations during the diplophonic phase (data not shown).

Figure 1 Normalized GAW derived from HSDI recordings of the MTD patient, showing a transition from (a), single-cyclic phase (phase I) to (b), diplophonic phase (phase II)

These findings suggest that it is the non-homogeneity in the A-P VF vibrations, but not the asymmetry, that is linked to the diplophonia. Interestingly, both asymmetry and non-homogeneity were evaluated to exist in the VF vibration prior to transitioning to the diplophonic phase, and is characterized by glottal incompetence and a breathy and rough VQ. Our analyses also revealed an improvement in both symmetry and homogeneity of the VF vibration after the phase transition (during diplophonic phase) (data not shown).

Figure 2 GAW Nyquist plots representing three consecutive time periods of HSDI recordings (0~300ms; 300~600ms and 600~900ms) before (upper row) and after (lower row) Tx
Results from an identical analysis of HSDI recordings from the same patient after voice Tx revealed a significant improvement in the synchronization of vibrations within the VF that led to almost normophonic GAW and Nyquist pattern (Fig. 2).

Further, the jitter measures were calculated and the results showed that post-Tx voice was significantly improved as evaluated by a jitter of 1.9%, as compared to 8.8% for pre-Tx voice during phase I, or 2.2% and 2.5% during phase II (diplophonic phase), representing perturbation measures for the two respective simultaneous vibrations.

IV. DISCUSSION

Quantitative evaluation of symmetry/homogeneity of VF vibrations in the pre- and post-Tx MTD voice were performed and the results showed an overall improvement in the VF vibration after Tx. In particular, the anterior (A), medial (M) and posterior (P) portions of the vocal folds vibrate at the same frequency (Fo~194 Hz) and the A-M-P vibrations are better synchronized, as evidenced by high correlation coefficient between the A-P vibrations (0.6171), compared to pre-Tx results (-0.13), indicating complete out-of-phase vibrations in the A-P vocal folds.

V. CONCLUSION

To conclude, asynchrony in VF vibration at different portions of the A-P axis of the vocal folds is associated with the diplophonia perceived in this MTD patient. Analyses of the HSDI recordings from the same patient after Tx demonstrated significant improvement in vibratory synchronization along the VF length, consistent with the subject having a normal perceptual rating after Tx as rated by the GRBS (G-grade, R-roughness, B-breathiness and S-strain) scale.

REFERENCES