



Review of Current Practices and Preliminary Recommendations for Instrumental Singing-Voice Diagnostics

Überprüfung der aktuellen Praxis und erste Empfehlungen für eine instrumentelle Diagnostik der Singstimme

Calvin Peter Baker^{1, 2, 3}

¹Speech Science, School of Psychology, University of Auckland, 1010, Auckland, New Zealand

²School of Music, University of Auckland, 1010, Auckland, New Zealand

³Swallowing and Voice Research Laboratory, University of Auckland, 1010, Auckland, New Zealand

Received 1 September 2025, accepted 7 November 2025

Abstract

Advances in acoustic and aerodynamic instrumentation now afford clinicians and researchers a wide range of voice analysis tools for diagnostics and quantitatively measuring treatment outcomes. Within clinical populations, singers are a unique population requiring specialised diagnostic approaches. Their voice complaints are often highly nuanced and context-specific, and in some cases, there are no readily observable structural or functional changes at the level of the vocal folds. To date, few works have explicitly focused on clinically relevant diagnostic approaches for singers. In the absence of laryngeal findings through visualisation, the singing voice should be dynamically assessed using a range of techniques to shed light on 'zones of interest' in the client's voice, related to their self-reported voice concern. Questions remain, however, as to what readily available, cost-effective tools can be successfully used with singing-voice samples in clinical settings. This narrative review discusses relevant acoustic and aerodynamic voice metrics that can be useful for quantifying singing-voice function across a range of sung tasks. Principles of vocal efficiency and economy are discussed, encouraging clinicians and medical professionals to consider the vocal toll of arbitrarily aiming to ameliorate inferential voice metrics (for example, reducing perturbation or increasing sound pressure level). Preliminary recommendations for singing-voice analysis are also offered, along with a protocol and PRAAT script to calculate a range of relevant acoustic voice measures for quantitative and qualitative singing voice analysis.

Abstrakt

Fortschritte in der akustischen und aerodynamischen Instrumentierung bieten Klinikern und Forschern heute eine breite Palette von Tools zur Stimmanalyse für die Diagnostik und quantitative Messung von Behandlungsergebnissen. Innerhalb der klinischen Populationen stellen Sänger eine einzigartige Gruppe dar, die spezielle diagnostische Ansätze erfordert. Ihre Stimmbeschwerden sind oft sehr nuanciert und kontextspezifisch, und in einigen Fällen sind keine leicht erkennbaren strukturellen oder funktionellen Veränderungen auf der Ebene der Stimmlippen vorhanden. Bislang gibt es nur wenige Arbeiten, die sich explizit mit klinisch relevanten diagnostischen Ansätzen für Sänger befassen. In Ermangelung visueller Befunde aus dem Kehlkopf sollte die Singstimme dynamisch mit einer Reihe von Techniken beurteilt werden, um „interessante Bereiche“ in der Stimme des Patienten aufzudecken, die mit seinen selbst berichteten Stimmproblemen in Zusammenhang stehen. Es bleibt jedoch die Frage, welche leicht verfügbaren, kostengünstigen Instrumente erfolgreich mit Singstimmproben in klinischen Umgebungen eingesetzt werden können. Diese narrative Übersicht diskutiert relevante akustische und aerodynamische Stimmmetriken, die für die Quantifizierung der Singstimmfunktion bei einer Reihe von Gesangsaufgaben nützlich sein können. Es werden Prinzipien der Stimmleistung und -ökonomie diskutiert, um Kliniker und Mediziner zu ermutigen, die stimmlichen Belastungen zu berücksichtigen, die mit dem willkürlichen Bestreben einhergehen, inferentielle Stimmmetriken zu verbessern (z. B. durch Verringerung von Störungen oder Erhöhung des Schalldruckpegels). Außerdem werden vorläufige Empfehlungen für die Analyse der Singstimme gegeben, zusammen mit einem Protokoll und einem PRAAT-Skript zur Berechnung einer Reihe relevanter akustischer Stimmparameter für die quantitative und qualitative Analyse der Singstimme.

Keywords:

voice diagnostics, singing voice, voice outcome measures, acoustic voice analysis, aerodynamic voice analysis

Schlüsselwörter:

Stimmdiagnostik, Singstimme, Leistungsbeurteilung der Stimme, akustischer Stimmanalyse, aerodynamischer Stimmanalyse.

INTRODUCTION

Singers are professional or vocational voice users who are at high risk for vocal injury and development of voice disorder (Childs et al., 2023; Mori et al., 2017; Pestana et al., 2017, 2023; Phyland, 2017; Phyland & Miles, 2019; Prebil et al., 2020). Compared with speech, sung phonation delivers increased intensity, frequency, and duration of phonation, which add to overall vocal demand (Calvache Mora et al., 2023). In combination with social and environmental factors associated with the performing arts – including poor air quality, vocalising in loud settings (Lombard effect), alcohol and substance use, and poor sleep hygiene – these demands further increase the likelihood of vocal fatigue, injury, or disorder (Calvache Mora et al., 2023; Hunter et al., 2020; Pestana et al., 2017). Although it is well known that singers face increased risk of voice disorder (Phyland, 2017; Phyland & Miles, 2019), relatively little work has been carried out specifically exploring clinical diagnostic techniques that are relevant and appropriate for singers.

In all instances, visual examination of the vocal folds is a recommended first step in clinical diagnosis of voice complaints (Lechien et al., 2023; Patel et al., 2018; Sataloff, 2017; Welham, 2009). Structural and, to some extent, functional findings (Powell et al., 2020) are often observable using a variety of imaging techniques, including transnasal videolaryngostroboscopy and high-speed videoendoscopy. High-speed video recording provides extremely detailed view of vocal-fold vibration and offers, in combination with videokymography, unique ways of observing the true glottal cycle (Dippold et al., 2015; Herbst et al., 2016; Patel et al., 2023; Yamauchi et al., 2016). However, access to such equipment is extremely limited, particularly within clinical and public health sectors. Visualisation of the vocal folds alone is not sufficient to obtain a well-rounded understanding of voice attributes including function and quality. Thus, when combined with other instrumental and perceptual measures such as those gathered using aerodynamic, acoustic, and self-report tools, a clearer picture can be formed of the nature and impact that a voice complaint has on a client's vocal function. The use of diagnostic tools from a variety of domains strengthens medical decision-making by supporting (or questioning) initial medical findings through visualisation, and provides useful quantitative metrics with which to measure treatment outcomes.

When no visual findings are present, it can be challenging to ascertain the nature of the voice complaint as well as its true effect on vocal function. Skilled singers are finely attuned to their instrument, often

having developed heightened proprioception of vocal sensations during phonation (Scotto Di Carlo, 1994). Thus, when they report changes in singing sensations beyond the scope of normal day-to-day fluctuation in the absence of clear aetiology, the voice healthcare team has the task of determining the nature of such dysfunction. Timely diagnosis is crucial for the professional singer, as any delay may further increase risk of acute vocal injury or the development of maladaptive behaviours, and precipitate adverse effects on livelihood, mental health, and standing within professional communities (Phyland, 2017; Phyland & Miles, 2019). It is also crucial to note that vocal demand response (that is, the singer's conscious or unconscious functional response to a vocal demand) is affected by both the condition of the vocal system (internal) as well as environmental (external) contexts (Calvache Mora et al., 2023; Hunter et al., 2020; Shembel & Nanjundeswaran, 2022). Both need to be considered when attempting to define the cause of a singer's voice complaint.

The presence or absence of laryngeal findings in visualisation of the vocal folds does not, in itself, confirm the presence or absence of dysfunction or deviance from a client's normal voice production (Sataloff et al., 2012; Van Houtte et al., 2010). Rather, this may reflect the limitations of current capabilities of visualisation techniques as well as practitioner experience and knowledge surrounding optimal singing-voice function (Sataloff, 1981). For singers presenting with voice complaints that are not readily quantifiable via traditional visualisation techniques (or in the absence of laryngeal findings), the question remains: what tools can clinicians use instrumentally to assess the singing voice for the purpose of identifying biomarkers of singing-voice dysfunction? Greater understanding regarding the challenges that instrumental singing-voice analysis presents and more inquiry into specific tools that are suitable for clinical singing-voice assessment is required.

Methods

The present narrative review reports on currently available methods, tools, and protocols for instrumental singing voice analysis in both clinical and subclinical contexts. The purpose of the review is to provide the reader with an overview of potential instrumental voice metrics for use in clinical singing-voice analysis, and to identify gaps in knowledge for further investigation.

Given the narrow scope of this work, no specific dates were applied to the search. A particular focus was placed on diagnostic tools in the acoustic, aerodynamic and perceptual domains that have been

designed specifically for or tested with singing-voice samples. Priority was also given to those that can be readily implemented in a clinical setting with relatively minimal cost. Only English language literature was reviewed. Scholarly peer-reviewed literature was sourced using Google Scholar, Scopus, and PubMed. Search keywords (and combinations of these) included 'singer', 'singing', 'singing analysis', 'voice diagnostics', 'acoustic diagnostics', 'aerodynamic diagnostics', 'acoustic voice analysis', 'aerodynamic voice analysis', 'singing voice perception', 'vocal health', and 'perceptual voice analysis'. Bibliographies and reference lists of sourced literature were also used to find further works that may not be well indexed online. Gathered literature was screened for relevance to the central research question, namely, whether they commented on acoustic, aerodynamic, perceptual, or self-report measures of the singing voice.

Works reporting on voice metrics useful for singing-voice analysis are discussed in narrative, alongside considerations that should be made when calculating their values and interpreting data. Where necessary, supplementary theoretical background is also offered to support the reader's understanding of the measures discussed. Based on the review findings, several pragmatic and relatively easily employed recommendations are offered for implementing acoustic singing-voice diagnostics in clinical settings.

ACOUSTIC ANALYSIS

Acoustic voice analysis is widely used in clinical and research settings to quantify the degree of a voice signal's deviance from stable production, and to track treatment outcomes. Acoustic analysis offers

a flexible, cost-effective tool for instrumental voice analysis, through which many quantitative measures related to vocal outputs and inferred vocal function can be produced (Brockmann-Bauser, 2024; Gunjawate et al., 2018; Murray & Woodnorth, 2020; Schultz & Vogel, 2022; Ternström, 2005; Titze, 1995). However, consensus is yet to be reached regarding the physiological correlates of some metrics, and whether all are suitable for use with singing-voice samples.

Perturbation and Periodicity

Voice perturbation measures are near-ubiquitously integrated into standardised voice assessment protocols and indices (Barsties v Latoszek et al., 2017; Lechien et al., 2023; Maryn et al., 2010; Maryn & Weenink, 2015; Patel et al., 2018; Wuyts et al., 2000). Many previously reported data from both healthy and disordered voices are available. Measures such as *jitter* and *shimmer* reflect frequency and amplitude perturbation of the voice signal within the time domain. Figure 1 illustrates the derivation of these parameters from an acoustic signal.

Challenges in the calculation of time-domain parameters arise when waveform properties are highly variable within the analysis window (Titze, 1995). Time-domain-based measurements require accurate fundamental frequency (f_0) tracking (that is, cycle-to-cycle delineation), and so are susceptible to confounding effects of volitional changes in f_0 (Dejonckere et al., 2012; Titze et al., 1987; Titze & Liang, 1993). Commonly used acoustic voice measures including jitter, shimmer, and harmonics-to-noise ratio are also influenced by f_0 .

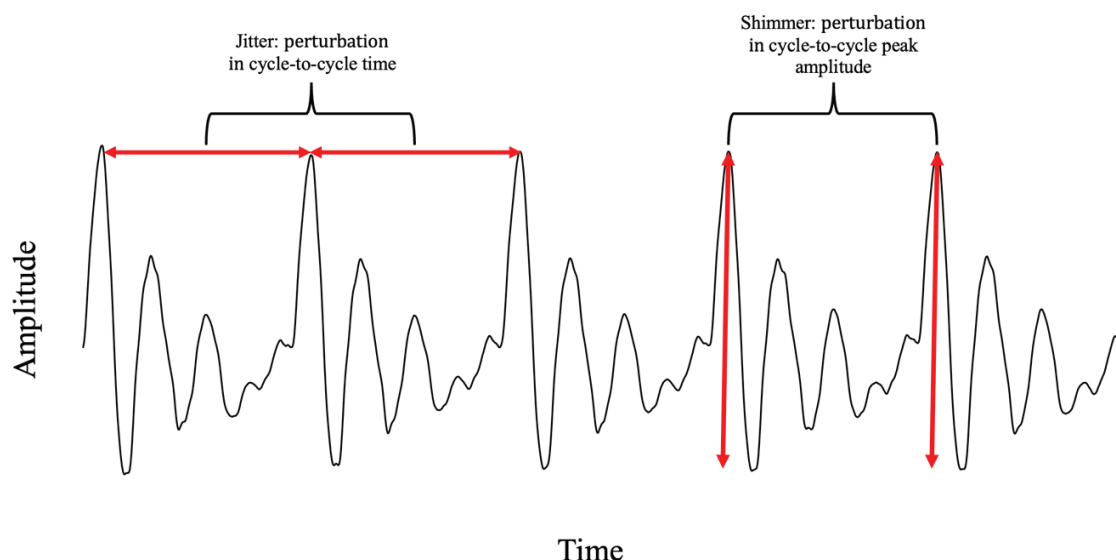


Figure 1. Illustration of Common Time-Domain Parameters Calculated from an Acoustic Voice Signal

and sound-pressure level (SPL), even within quasi-periodic normophonic voices (Brockmann et al., 2008, 2011; Brockmann-Bauser et al., 2018). These effects appear not only to be due to changes in vocal function. A recent in-silico (computer-based) study also demonstrated changes in these parameters in controlled synthesised tones (Baker, Brockmann-Bauser, et al., 2023) suggesting that acoustic resonance-harmonics interactions can also affect acoustic perturbation and noise measures. Singing-specific features such as vibrato have also been seen to systematically affect jitter, shimmer, and related metrics including relative average perturbation and amplitude perturbation quotient (Baker, Brockmann-Bauser, et al., 2023; Baker et al., 2024). Previous works have found similar deleterious effects of vocal tremor on perturbation measures to those of singing vibrato (Gillivan-Murphy et al., 2019; Ramig & Shipp, 1987).

More recently, cepstrum-based measures such as *smoothed cepstral peak prominence (CPPS)* have been utilised in voice research. The *cepstrum* is an inverse fast Fourier transform of the natural logarithm frequency spectrum, converted back to the time domain (for a thorough discussion on the cepstrum and its interpretation, the reader is directed to Fraile and Godino-Llorente [2014]). When a regression line is applied to the cepstrum, the decibel (dB) difference between the regression line and the most prominent peak in the cepstrum (usually representing f_0 in milliseconds) provides the cepstral peak prominence (CPP). If this calculation is made after applying a smoothing factor to the cepstrum, CPPS can be calculated. CPPS has demonstrated statistically significant associations with overall grade of dysphonia, as well as roughness and breathiness in both, sustained vowel and continuous speech samples (Hassan et al., 2021; Heman-Ackah et al., 2003; Hillenbrand et al., 1994; Hillenbrand & Houde, 1996; Murton et al., 2020).

CPP(S) shows promise for singing voice analysis as it does not require accurate f_0 tracking. However, in both healthy and disordered voices, CPPS is positively associated with f_0 and SPL (Brockmann-Bauser et al., 2021; Sampaio et al., 2020, 2021). Baker et al. (2024) used controlled synthesised tones to systematically and independently explore the effects of f_0 , source-spectrum tilt, and vibrato extent on CPPS. A non-monotonic relationship between CPPS and f_0 was found, where CPPS increased when the f_0 approached approximately 175 Hz, and then decreased again with higher f_0 -values. Across a range of -12 dB, source-spectrum tilt (-6 to -18 dB/octave) had a positive near-linear relationship with CPPS ($R^2 = .998$). Vibrato extent (here meas-

ured in semitones, the extent of f_0 deviation from the mean f_0) also influenced CPPS, where larger vibrato extents were associated with lower CPPS values, albeit to a small degree (approximately -2 dB change from 0.0 to 1.0 semitone vibrato extent). Given the highly systematic effects, the influence of f_0 and SPL could be successfully controlled for post hoc statically or by using a correction factor.

To detect subtle changes in singing voice function, the singing voice should be analysed, not solely industry standard comfortable pitch and volume sustained vowels or reading passages. Voice problems in singers may pertain to a certain frequency range or within specific musical or phonological contexts. The habitual speaking voice may in fact present no challenge for the singer and be perceptually normal if not excellent. Furthermore, the singer should be able to provide voice samples that are as ecologically valid as possible, that is, including all usual features of their singing voice where the problem may occur. These may include vibrato, formant tuning, formant clustering, and stylistic elements such as slides or changes in phonation type. Few works have specifically accounted for such variables in their analysis of the singing voice, leading to varied results across studies. Furthermore, f_0 , SPL, and vibrato characteristics are not routinely reported, hindering accurate data interpretation.

One particularly salient feature of the singing voice is the presence of *vibrato*. Although not ubiquitous across all styles, it is a feature of many and is generally considered a sign of a free, healthy voice production (Bonjyotsna & Bhuyan, 2016; Mitchell & Kenny, 2010; Nestorova et al., 2023; Nix et al., 2016). Vibrato has been described as a quasi-periodic oscillation of f_0 at a certain rate (Hz) and extent (frequency deviation from the mean or target f_0 often measured in cents or semitones) and also includes periodic changes in airflow rate and resonance-harmonics interactions (Horii, 1989a, 1989b; Nandamudi, 2017; Shipp et al., 1990; Sundberg, 1995). Given the main driving component of vibrato is related directly to voice source behaviours (cyclical changes in vocal-fold longitudinal tension), assessing its regularity is likely to offer useful information in voice analysis and potentially as an early biomarker for singing-voice dysfunction.

BioVoice (Morelli et al., 2021) is a MATLAB built software designed to accurately analyse voice signals, including high frequency and quasi-stationary singing samples. The algorithms developed for f_0 detection (Sangiorgi et al., 2009) have shown high reliability in noisy environments and appear to outperform other commonly used techniques (Manfredi et al., 2012; Orlandi et al., 2013; Rruqja et al.,

2014). Within BioVoice, singing-specific parameters including vibrato rate and extent can be calculated, as well as vibrato jitter (V_{Jitt}) and vibrato shimmer (V_{Shim}). Similar to their traditional counterparts, V_{Jitt} and V_{Shim} quantify the amount of perturbation in the rate and extent of the vibrato f_0 oscillation. Manfredi et al. (2015) used these measures when comparing differences between operatic and jazz singing styles and found that operatic samples were characterised by a more regular vibrato (lower V_{Jitt} and V_{Shim}). Baker (2023) and Baker et al. (2023a) found that V_{Jitt} and V_{Shim} were significantly and positively associated with the vocal fatigue subscale of the Evaluation of Ability to Sing Easily (EASE, discussed further below; Phyland et al., 2013, 2014). These findings suggest the utility of vibrato-perturbation based measures for singing voice analysis, and points to the possible use of vibrato perturbation as a biomarker of early singing-voice dysfunction. Recent advancements have also been made in the analysis of vocal tremor characteristics (Maryn et al., 2019, 2025). Given the similarities between vocal tremor and singing-voice vibrato (Gillivan-Murphy et al., 2019; Ramig & Shipp, 1987; Titze et al., 2002), it is possible that these tools could be extended and applied to the singing voice in future work. Qualitative inspection of vibrato characteristics is a simple way to observe a voice-source related parameter and can provide insights into singing-voice stability.

The Spectrum and Spectrography

The *spectrum* and *spectrogram* of an acoustic signal provide a useful means of visualising the harmonic and resonance components of the voice, offering unique yet complementary data to time-based measures. Usually, fast Fourier transform is used to deconvolve a complex time-domain waveform (for example, an acoustic voice signal) into its individual frequency components. These are windowed and represented in a plot that provides frequency on the abscissa and amplitude on the ordinate in the case of a spectrum, and time and frequency on the abscissa and ordinate, respectively, in the spectrogram. The spectrogram includes the third parameter of relative intensity of the frequency components, which are usually represented in vividness or shade of colour. Figure 2 shows an example of a spectrum (panel A) and a spectrogram (panel B) of a sustained /a/ vowel in a male speaker.

Using the spectrum, relative dB differences between harmonic components (and harmonic and noise components) can be measured to provide data related to phonation type (*breathy* to *pressed*, correlated with glottal closure characteristics), vocal tract resonances,

and harmonics-to-noise ratios (Childers & Lee, 1991; Cleveland & Sundberg, 1988; Gramming & Sundberg, 1988; Hammarberg et al., 1980; Hejná et al., 2019; Lă et al., 2023; Sundberg, 1974, 2001, 2022).

Another relevant feature of the spectrum is *formant clusters*. Specifically, the so-called *singer's formant cluster (SFC)* is a gathering of the third, fourth, and fifth vocal tract resonances within approximately 2-4 kHz (Sundberg, 1974, 1987, 2001; Sundberg et al., 1993, 2013). The SFC is perceptually relevant, lending 'brightness' and 'ring' to the sound, and allows unaccompanied singers to be heard over acoustic instruments and orchestras (McCoy, 2012; R. Miller, 1996). Lă et al. (2023) used the *Formant Cluster Prominence (FCP)* to quantify differences in the singer's formant cluster in Fado singers. The FCP is calculated by first adding a linear trendline in the 1-5 kHz region of the spectrum to account for the effect of SPL on spectral tilt (Nordenberg & Sundberg, 2004; Sundberg & Nordenberg, 2006). The dB difference between the spectrum peak in the 2-4 kHz region and the trendline at the same frequency is then measured, providing the FCP. As SPL is somewhat accounted for, the measure can be compared between subjects and measurement points. The FCP may be a useful quantitative metric for measuring resonance-related characteristics of the singing voice. It should be noted, however, that the SFC is not always present when phonating at

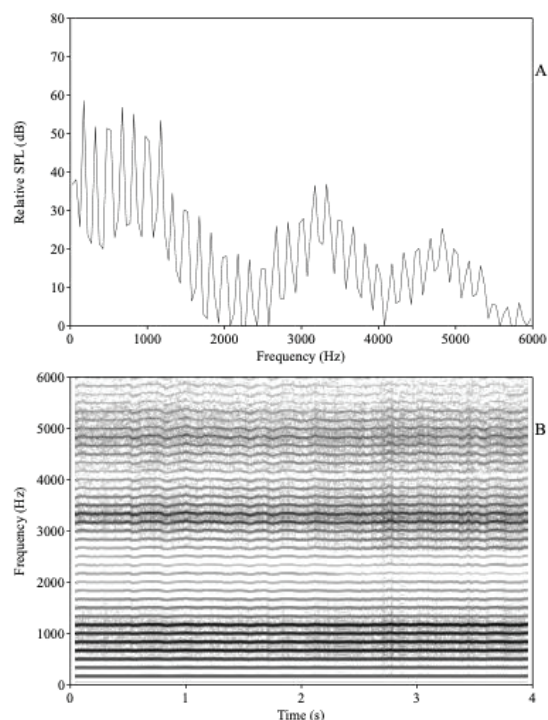


Figure 2. Spectrum (A) and Narrowband (50Hz) Spectrogram (B) of a Spoken Sustained /a/ Vowel

higher frequencies, as is the case in higher voice types such as sopranos (Weiss et al., 2001). Thus, voice type and sample f_0 should be considered when interpreting FCP data. The FCP can be easily calculated in the spectrum module of the freeware *Sopran* (Granqvist, 2025c).

The spectrogram offers a slightly more intuitive view of the acoustic signal which provides opportunity for observation and qualitative analysis of a sung voiced sample. As the voice is reflected over time, tasks such as glides throughout the singer's range offers insights into areas of instability, for example when crossing a register transition (Behlau et al., 2023; Kob et al., 2011). The spectrogram also provides a simple way to visualise a singer's vibrato, which can provide useful information related to singing-voice stability and regularity. As visualised in the spectrogram, a stable vibrato should be represented by a quasi-sinusoidal oscillation in f_0 , which can be easily observed in upper harmonics. The spectrogram can also provide a simple but effective tool for visualising *bifurcation* or *subharmonics* in a voice signal, indicating deviation from stable and periodic vibration of solely the true vocal folds (Omori et al., 1997; Titze, 2000). These techniques have also been incorporated into pedagogical practice for decades and are currently used as biofeedback tools for voice training and speech-language therapy (Baken & Orlikoff, 2000; Bozeman, 2013; Lã & Fiuza, 2022). Echter-nach and colleagues (2017) used spectrography alongside high-speed video endoscopy and electroglottography to evaluate pre-post treatment effects in singers across the passaggio (register transition region). In their work, the effects of vocal fold mass lesions on participants' vocal stability across the passaggio region could be clearly seen in spectrograms.

Figure 3 shows spectrograms of a singer performing an ascending and descending pitch glide (panel A) and a softly sung /a/ vowel (panel B). Note the clear registrational 'breaks' (A) and the somewhat unstable vibrato and the presence of occasional subharmonics in the region of 2.5 to 4 seconds (B). Behlau et al. (2023) proposed dynamic voice analysis using spectrography to inspect vocal function across a variety of voiced tasks, each of which has clear causal physiological links (for example, cricothyroid contraction, vocal fold lengthening, adductory quality, and vibratory stability). A similar protocol for the singing voice is discussed later in this work.

Voice Range Profiles

One elegant solution to the aforementioned covariances between f_0 , SPL, and acoustic voice measures is to map the voice across f_0 and SPL. The

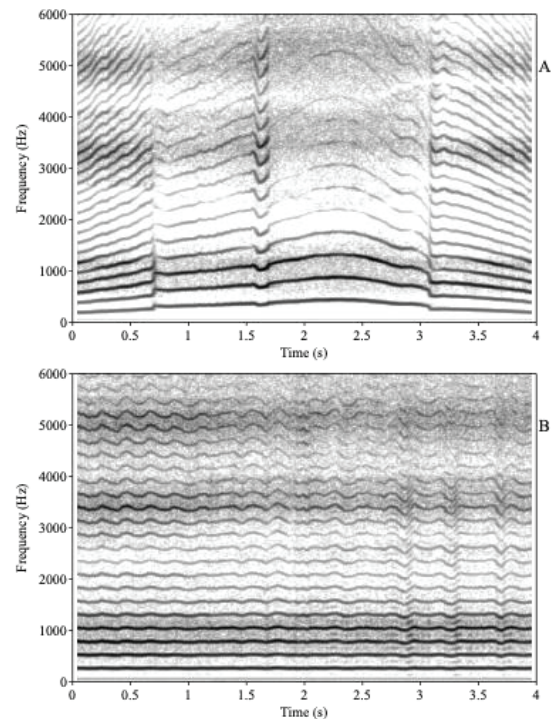


Figure 3. Spectrograms of a Singer Performing a Pitch Glide with Registrational Breaks (A), and Unstable Vibrato and Sub-Harmonics on a Sustained C4 (B)

resulting *voice range profile* (VRP) provides a picture of a client's functional frequency and SPL range. Gaps in the voice can be observed, as well as areas where f_0 or SPL are compromised. Pre- and post-treatment comparisons of VRPs can provide useful insights into treatment outcomes. If a singer has a baseline VRP of healthy phonation, this can also be used to detect changes in the voice that may be difficult to perceptually assess in real-time. Normative data regarding VRP characteristics for singers is currently sparse and comprised of comparatively small samples (Lycke & Siupsinskiene, 2016; Pabon et al., 2014).

Lamarche et al. (2009) proposed the collection of traditional VRP data with the addition of participants' perception of vocal instability. When performing the VRP protocol, participants pressed a trigger which recorded binary pulses synchronised to the acoustic voice data. When plotted against the VRP the result showed areas where participants felt the loss of vocal control or stability. As expected, there were more perceptual responses at the extremes of the VRP, that is, when singing at lowest and highest f_0 s and SPLs. When exploring differences between healthy and disordered singers, Lamarche, Mor-somme, et al. (2009) observed group differences in the locations of reported vocal instability: healthy singers tended to report instability only at vocal

extremes, whereas disordered singers reported instability in a narrower f_0 and SPL range.

The inclusion of singer perception is highly relevant to the clinical assessment of singing voice complaints, particularly in the absence of clear diagnosis based on visualisation. As previously discussed, professional singers are attuned their instrument and able to recognise deviation of function from their baseline norms. By simultaneously incorporating self-report with instrumental voice measures into a clinical routine, areas of instability in a singer's f_0 and SPL range may be better illuminated for the clinician and thus help guide treatment decisions.

Works by Pabon, Ternström, Lamarche and colleagues have advanced the traditional VRP (or phonetogram) to so-called *voice mapping* to include additional voice metrics such as electroglottographically (EGG)-derived contact quotients, and acoustic perturbation, periodicity, and spectrum-based measures (Lamarche et al., 2010; Pabon, 1991; Pabon et al., 2011, 2014; Pabon & Ternström, 2020; Ternström & Pabon, 2022). The VRP and voice mapping have been successfully used to describe salient singing voice parameters such as f_0 and SPL range, voice-source characteristics across the range, and training effects (Lamarche et al., 2010; Pabon et al., 2014).

Challenges in implementing VRPs and voice mapping in clinical contexts include the time required to elicit quality data, potential vocal fatigue and reduction in voice quality due to inherent voice demands (phonation at vocal extremes), as well as limited access to user-friendly software for clinical use. Nonetheless, the inclusion of additional voice metrics makes VRPs and voice mapping a useful addition to the toolkit for singing-voice diagnostics. One easy-to-install VRP/voice mapping programme is *RecVox* (Granqvist, 2025b). The stand-alone freeware provides a relatively simple-to-use interface that can record and display a VRP in real time, as well as calculate a third parameter, including accumulative time (seconds); L_1 - L_2 (dB); harmonics-to-noise ratio (HNR, dB); the spectral power centroid (Hz); crest factor (dB); maximum partial number; and f_0 standard deviation (cents). A more detailed discussion of these parameters and the software's technical details can be found in Granqvist (2025a). Of particular interest is L_1 - L_2 measure (the dB difference between the first and second harmonics, also termed H_1 - H_2), which correlates with the glottal flow pulse amplitude and relates to phonation type from breathy to pressed (Gauffin & Sundberg, 1989; Sundberg, 2022). The L_1 - L_2 measure offers some easily attainable, physiologically-linked data on voice production in the singing voice. Although spectral measures are generally

affected by SPL and f_0 (Nordenberg & Sundberg, 2004; Sundberg & Nordenberg, 2006), these are somewhat automatically accounted for by the nature of the data mapping f_0 and SPL, showing these associations clearly. Another freeware available for VRP and voice mapping is *FonaDyn* (Ternström et al., 2018). This software incorporates many of the above-mentioned acoustic measures, as well as having the additional functionality of including CPP and electroglottographically derived measures. For a thorough review of the VRP, phonetogram, and voice mapping, the reader is directed to Ternström and Pabon (2022). Future research should explore the diagnostic utility of VRPs and voice mapping for use with singers and their association with other perceptual measures (discussed below).

Acoustic Indices

Although there is a plethora of acoustic voice metrics available for use with relative ease, no single value can holistically capture the nuances of voice production. Diagnostic power and associations with perceptual voice evaluation increase when individual voice metrics are combined into an index (Barsties v. Latoszek et al., 2025). Examples of these for speech include the *Dysphonia Severity Index* (Wuyts et al., 2000), the *Acoustic Voice Quality Index* (Maryn et al., 2010, 2013; Maryn & Weenink, 2015), and the *Acoustic Breathiness Index* (Barsties v Latoszek et al., 2017).

To date, the so-called *Aeroacoustic Singing Fatigue Index* (ASFI; Baker, 2023; Baker, Purdy, et al., 2023a) represents the only proposed multidimensional index constructed specifically with clinical singing voice analysis in mind. The ASFI utilises both acoustic and aerodynamic voice metrics including CPPS, V_{Jitt} , V_{Shim} , and an efficiency ratio (the ratio between SPL and subglottal pressure [P_{Sub}]). These are calculated from a softly sung /a/ vowel on C4 or C5, and a softly sung consonant-vowel train /papa-papapa/ on a C4 or C5 for lower or higher voices, respectively. In a cohort of 18 healthy classical singers, the index showed significant association with the *Evaluation of Ability to Sing Easily* (EASE, Phyland et al., 2013, 2014) vocal fatigue subscale ($p < .001$), and was able to explain 62.4% variance in the composite score of three vocal fatigue subscale questions: 'My voice is husky', 'My voice is dry/scratchy', and 'My top notes are breathy'. Although somewhat cumbersome to calculate, the concept shows promise. As a singer's perception of their own vocal effort and condition is likely related to several different sensation-based parameters (for example, pressure, tension, and vibration), the ASFI includes measures

related to both, the vocal output as well as the input, that is, aerodynamic effort. Breathiness in the singing voice (when not an artistic choice) can reflect glottal incompetence and reduces CPPS. A singer's vibrato regularity and ease are readily perceived and are related to V_{Jitt} and V_{Shim} . The input necessary to generate a softly sustained C4 or C5 may also give a singer clues as to the condition of their voice. This is reflected in the efficiency ratio, where higher pressures relate to reduced ease of phonation.

Much more testing and refinement of the ASFI is needed. It would be useful to incorporate data collection within one software programme, and to explore the index's performance in more diverse cohorts of both healthy and disordered singers. The use of a simple *intra-oral pressure metre* and *omni-directional headset condenser microphone* may suffice if recorded in synchrony and properly calibrated. The inclusion of a *pressure-based metric* and its significant association with singers' perception of their vocal condition, specifically related to vocal fatigue, is highly relevant. While acoustic measurement of the voice provides useful information related to the vocal *output*, pressure values provide information on vocal *input*.

AERODYNAMIC ANALYSES

While most clinical voice measures are derived from the vocal output (that is, the radiated acoustic signal), there are relatively few that consider vocal input, that is, pressure and flow. This is no doubt related to the lack of access to instrumentation that reliably measures aerodynamic voice properties in a clinical setting and at a feasible cost for public health. *The Phonatory Aerodynamic System* (PAS, KayPentax, n.d.) is a handheld transducer system connected to a facemask that collects measurements of airflow (L/s), volume (L), and intra-oral pressure (cm H₂O) during phonation. These can be used to infer glottal behaviour during a wide range of phonatory tasks.

Breath management is a fundamental principle within voice pedagogy, and efficiency of breath use is often discussed and promoted (Baker et al., 2025; Herbst, 2017). In a clinical setting the ability to measure airflow during sung phonation provides clinicians with invaluable insights into glottal valving behaviours. Higher airflow rates may suggest reduced glottal completeness, relative to f_0 and SPL (Allen, 2019; Bouhuys et al., 1968; Iwata et al., 1972). Flow rates may be measured pre- and post-treatment during controlled sung tasks to assess changes in glottal completeness. A device similar to the PAS is the *Aeroview System* by Glottal Enterprises (Syracuse, NY, USA), which, unlike the PAS, provides an acoustic

signal that is also suitable for acoustic analysis and inverse filtering (Rothenberg, 1973). The Aeroview System was designed for research use and has not yet been validated for clinical applications.

The cost of commercial devices for aerodynamic voice measurement may be generally prohibitive in public healthcare sectors. However, mean flow rate can be calculated using a simple differential equation (Equation 1) by dividing the *expired air volume* (which could be collected using a relatively less expensive handheld spirometer) by the *phonation time*. Thus, even in contexts where sophisticated aerodynamic analysis is not possible, some insight can be gained into averaged flow rates. These can be compared pre-/post-treatment, as well as between healthy baseline and disordered measurements, provided the task instructions remain the same.

$$\text{Mean flow (L / s)} \approx \frac{\text{Expiratory Volume (L)}}{\text{Phonation Times (s)}} \quad (1)$$

Intra-oral pressure measured during a 'p' occlusion within a /papapapapa.../ task provides an inferential measure of P_{Sub} (Hertegård et al., 1995; Smitheran & Hixon, 1981). Baker et al. (2023a) found a significant correlation between singers' self-perceived voice condition and the ratio between SPL and P_{Sub} . Measurement of pressure events in the context of sung phonation may provide invaluable insights into the amount of effort singers use to generate voiced sound, and may be related to their perception of vocal fatigue or discomfort. Pressures before and during phonation may also provide insights into hypo- or hyper-functional voice behaviours (Herbst et al., 2015; Toles et al., 2022). Figure 4 illustrates the relationship between pressure and flow in neutral and pressed phonation.

Another pressure-based measurement likely related to the perception of vocal effort is the *phonation threshold pressure (PTP)*, that is, the pressure required to initiate vocal fold vibration for phonation (Plexico et al., 2011; Titze, 1992). Enflo et al. (2013) point out that phonation may occur without vocal fold collision (for example, in some very breathy phonation) and suggest that *collision threshold pressure* may be more appropriate (the pressure required to initiate vocal fold collision during phonation). Whether the vocal folds contact during a vibratory glottal cycle is also related to adductory forces and requires confirmation of vocal fold contact for its calculation (for example, through visualisation or electroglottography).

Vocal load increases PTP (Chang & Karnell, 2004; Enflo et al., 2013). Titze et al. (1995) found a link between fluid viscosity and PTP, where higher vis-

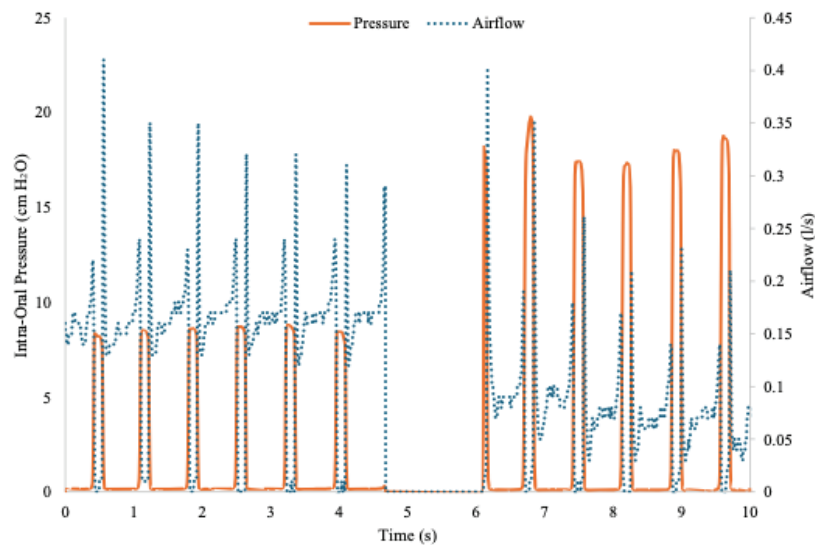


Figure 4. Relationship Between Pressure and Flow in Neutral (Left) and Pressed (Right) Type Phonation

cosity in modelled vocal fold epithelium required higher pressures to induce oscillation. Those with vocal fold oedema, structural alterations such as sulci or space-occupying mass, or hyperfunctional behaviours (whether primary or secondary) may demonstrate increased PTPs (Hillman et al., 2020; Plexico et al., 2011).

In singers, voice condition can be self-monitored via the *perception of the ease of production*. Within the context of a singer's healthy baseline, increased effort of phonation may be an early biomarker of voice dysfunction that is otherwise challenging to visualise or measure in the vocal output. Conversely, measuring the amount of input effort required to *initiate phonation* (that is, the PTP) may provide quantitative corroboration for self-reported increases in vocal effort. Naturally, PTP is also related to f_0 – relatively greater longitudinal tensions in the vocal folds require higher P_{Sub} to initiate oscillation of the vocal fold cover. Thus, any f_0 effect should be accounted for and reported alongside the data. Aerodynamic measurements provide invaluable insights into voice production input, its association with perceived effort and discomfort during singing, and as a biomarker of early voice dysfunction. Used in concert with acoustic and self-report measures, a clinician can obtain a range of data related to perceptual, acoustic (output), and physiological aspects of singing voice production.

AUDITORY-PERCEPTUAL AND SELF-REPORTING TOOLS

The *auditory-perceptual element* of the singing voice is a crucial component to consider when working

with singers in a clinical context and should be taken into account along-side self-reports and instrumental measures of vocal function. While a wide range of auditory-perceptual and self-report instruments are available for voice assessment, few have been designed specifically for singers.

Oates et al. (2006) developed an auditory-perceptual scale for the assessment of the classical singing voice. The items included overall vocal performance; appropriate vibrato; resonance balance; ring; pitch accuracy; breath management; evenness throughout the range; and strain. These offer a wide range of perceptual metrics with which to assess the singing voice. The scale was specifically designed for assessment of the classical singing voice, and as such may be less applicable to some contemporary styles where characteristics such as vibrato, ring, and evenness are not necessarily prioritised (Acosta Martínez & Daffern, 2023; Barlow & LoVetri, 2010; Butte et al., 2009; Sundberg et al., 1993; Thalen & Sundberg, 2001).

To measure the severity of voice disorder specifically in singers, the *Singing-Voice Handicap Index* (SVHI; Cohen et al., 2007) and the later SVHI-10 (Cohen et al., 2009) were created. The SVHI-10 provides a list of statements related to voice complaints experienced by singers and these are rated by the client on a scale from 0 (*Never*) to 4 (*Always*). The SVHI(-10) has shown ability to discriminate between healthy and disordered singing voices, providing a useful screening tool for singing voice disorder (in connection with vocal fold visualisation), and for tracking treatment outcomes. Perrin et al. (2025) found that the SVHI-10 had reasonable diagnostic accuracy ($AUC = .700$, $p = .003$) and

reported a minimal clinically important difference of ≥ 9.5 . The SHVI(-10) assumes a level of disorder or medical condition that may not be present in professional singers with more subtle voice complaints. As such, it may be less sensitive to preclinical changes in singing voice function.

To address this constraint, Phyland and colleagues (Phyland, 2014; Phyland et al., 2013, 2014) proposed the *Evaluation of Ability to Sing Easily* (EASE). This 22-question tool is made up of three subscales: Vocal Fatigue (VF), Pathological Risk Indicators (PRI), and Voice Concerns (VC). The VF and PRI subscales relate to sensation and perception-based elements of voice production (for example, effort, dryness, quality). The VC subscale predominantly reflects a psychometric measure of the degree of concern or worry about a singer's own voice. Each subscale should be scored separately and provides complementary but unique data on a singer's self-perception of their vocal health status, function, and voice complaint. The EASE was not constructed assuming the presence of voice injury or disorder, so it is particularly germane in the assessment of otherwise perceptually healthy singers. The EASE can be used regularly by the singer client to track their own voice condition during periods of high vocal demand (for example, long rehearsal or show periods), as well as to measure treatment outcomes.

To date, the EASE scale represents a unique tool that is not easily compared to other self-report instruments. A few recent studies found moderate correlations between the EASE and other validated scales, for example the Vocal Fatigue Index and translations of the SVHI-10 (Devin, 2023; D'haeseleer et al., 2025; Öğülmüş Uysal et al., 2024). Although further work to establish concurrent validity of the EASE and its subscales to related tools such as the Vocal Fatigue Index is of interest, they likely measure different aspects of fatigue at different sensitivities – trained singers may be more sensitive to symptoms of fatigue or experience these differently from untrained speakers. When using the EASE together with auditory perceptual ratings and the SVHI-10 for vocal health screening, the clinician is able to achieve a well-rounded view of (self)perception related to a singer's voice condition and function. Consideration must also be given to potential discrepancies between instrumental metrics and self-reported vocal effort.

EFFICIENCY VS ECONOMY – CONSIDERING THE TOLL OF THE OUTPUT

The term *efficiency* is widely used within voice pedagogy and therapy as an ideal characteristic of voice

production. Simply (albeit somewhat reductionistically), efficiency can be described as the absence of losses within a system when converting one type of energy into another. Within the voice system, this is the conversion of aerodynamic energy into acoustic sound energy, where the vocal folds act as a transformer. Intuitively, high efficiency is a good thing. Losses within the energy conversion process result in the dispersion of such energy which can contribute to the degradation of the system (for example, through heat accumulation within the vocal folds, friction, collision forces, as well as the addition of noise components into the radiated signal). Thus, it may be simple to state that efficiency should be a main priority when working with a singer with a voice complaint. Challenges to this assumption arise, however, when accounting for the potential physiological cost of arbitrarily measured efficiency.

Consider the example of increased SPL. A primary physiological correlate of SPL is P_{Sub} – higher P_{Sub} values are generally associated with higher SPLs (Baker et al., 2025; Ladefoged & McKinney, 1963). Increases in vocal fold closing-phase speed and contact phase duration are also associated with increased SPL (Alku et al., 2002; DeJonckere & Lebacqz, 2022; Ladefoged & McKinney, 1963; Stevens, 2000; Sundberg et al., 1999, 2005; Titze & Sundberg, 1992). This increase in acoustic output may reflect an increase in efficiency of the system – less wastage of aerodynamic energy in the conversion process. Indeed, as previously discussed, several widely-applied voice measures including jitter, shimmer, HNR, and CPPS all improve with increased SPL (Baker, Brockmann-Bauser, et al., 2023; Brockmann et al., 2008; Brockmann-Bauser et al., 2018, 2021; Orlikoff & Kahane, 1991; Pabon, 1991; Sampaio et al., 2021). We may reason that the system is always healthier when it is more efficient due to reduced losses.

Economy in a system can be described as the minimum expenditure for the maximum gain. In the voice system, this may be translated to the minimum aerodynamic and muscular input expenditure for the maximal acoustic output. Simply, how much pressure and muscular force is needed to phonate at a desired f_0 , SPL, and quality. While this value will change depending on the target output, the principle of economy asks for only the minimum effort to achieve the desired output. This would not preclude concurrent efficiency in most cases, as a clear, healthy sounding voice at an adequate SPL and without excessive breathiness (that is adequate glottal completeness) is an output goal for most singers (even if only in a baseline, neutral [not stylised] type phonation). The opposite is not necessarily the

case – the most efficient sound, that is the setup where losses are low at the level of the vocal folds, will not necessarily also be the most economical sound. *Pressed* phonation, for example, is usually characterised by high P_{Sub} , low flow, and increased SPL (acoustic output; recall Figure 4).

When working with singers, clinicians must consider not only the output of the sound, but also the input required for such and its potential toll on the system. Voice measures that suggest improved vocal function simply by virtue of increased phonatory efficiency (here described as reduced losses), should be interpreted with caution. Increased CPPS or HNR may result from a louder sound, but this may not reflect economy within the system and may lead to greater risk of vocal fatigue or injury. In fact, a highly pressed sound, if it is also free of subharmonics, may produce high CPPS and HNR values.

When considering the phonatory characteristics of a singer and how these are related to instrumental voice measures, several questions should be considered. *How can we assess subjective vocal effort in singers? How much effort is being exerted for the sake of ‘improved’ vocal metrics? Does the quantitatively measured output match the singer’s perception of effort or vocal comfort? What does the singer’s vocal output and input reveal about the function of the system as a whole?* More focus should be placed on the sensations of singing, vocal effort, and singer perception to lead the diagnostic inquiry, data interpretation, and treatment decision making.

SUMMARY, MAIN CRITIQUES, AND FUTURE DIRECTIONS

Voice clinicians and researchers are fortunate to have many instrumental tools to choose from when considering speaking voice assessment and diagnostics. Where the singing voice is concerned, the scope of appropriate measurement techniques is somewhat smaller. Within the acoustic domain, measures such as jitter and shimmer are widely applied but are influenced by large f_0 , SPL, and signal-quality effects when not controlled properly. Recording environment, microphone type, microphone-to-mouth distance, and signal-to-noise ratio should also be considered when using time-domain perturbation measures. Cepstrum based measures are more robust to these effects, but are also influenced by SPL, f_0 , and vibrato characteristics. As these effects are largely linear, their influence can be accounted for *post hoc* if needed. Spectrum and spectrography-based measures can non-invasively offer inferences on both, voice source behaviours

as well as resonance characteristics during a sung task. Perturbation in a singer’s vibrato can be quantified using a programme such as BioVoice, and qualitatively assessed in a spectrogram. Multiparametric acoustic indices may be useful additions to the diagnostic toolkit, but more work is needed to establish singing-specific tools. A VRP or voice map provides a baseline overview of the singing voice and vocal function that can be compared pre-/post-treatment.

In the aerodynamic domain, flow and pressure measurement offers invaluable insights into the effort used during habitual singing, as well as in prescribed tasks. The PTP reveals how much pressure is required to initiate phonation and, when inflated, suggests reduced vocal fold motility. Soft phonation tasks in general are likely to reveal vocal challenges more readily than habitual or loud phonation where vocal instability may be masked. Variability in airflow during sustained phonation may also indicate vocal fold vibratory irregularity and compromised laryngeal valving action.

Auditory perceptual and self-report tools designed with the singing voice in mind can provide clinicians with invaluable insights into a singer’s perception of their voice condition. Considering the singer’s perception of their own voice, as the expert, should be a standard part of any singing-voice assessment protocol. In particular, the SVHI-10 may be useful for tracking pre-/post-treatment in a disordered voice, and the EASE to track changes in voice perception related to vocal demand.

Across all measures, recording, reporting, and accounting for f_0 and SPL is crucial. Without such information, it cannot be determined whether changes in any one metric is a treatment effect or an artefact. Clinicians should also take care to properly calibrate acoustic and aerodynamic recording equipment to ensure comparability between measurement points. Much more work is needed to 1) determine which voice metrics can be reliably used with singing voice samples, 2) understand which tasks elicit the behaviours that reveal true voice condition in singers, and 3) construct a corpora of normative data that can guide clinicians in determining when a change in any one metric may be an indicator of singing-voice dysfunction. Furthermore, special consideration must be given to the challenges in understanding the difference between disordered singing and artistic choices. While free and regular vocal fold vibration should be a priority in any singer’s baseline phonation, there is much scope for variation from the ‘norm’. These must be carefully weighed up in view of vocal health, efficiency, and economy whilst ensuring that the professional sing-

ers' vocal identity is not undermined. Future studies considering the relationships between self-reported voice condition and instrumental voice metrics may contribute to identifying biomarkers of early singing-voice dysfunction that are crucial for timely and accurate preventative voice care.

PRELIMINARY RECOMMENDATIONS FOR CLINICAL SINGING VOICE DIAGNOSTICS

Foremost, the absence of clear laryngeal findings does not automatically indicate hypochondriasis. As singers are highly attuned to their vocal system and small deviations from its baseline, their opinions and comments should be carefully considered when assessing the singing voice and used to guide investigations. Thus, self-report measures such as the *SVHI-10* or *EASE* should be used at each clinical visit, so as to track the singer's own perception of their voice, vocal function, and phonatory effort. Considering the individual questions as well as the scales as a whole will provide clues to the cause of the problem.

When deciding which tools should be used in a battery for instrumental voice diagnostics, a clinician should first determine the physiological process they are seeking to find insights into. If there is suspected instability of the voice, the use of spectrography, perturbation measures on selected frequencies, and airflow variability measures may be useful to determine the scope and locus of such instability. If local hyperfunction is suspected, assessing PTP, mean pressures using a /papapa.../ task, and spectrum-based measures such as L_1 - L_2 may be useful to determine the severity of the condition on vocal function and track treatment effects. A VRP or voice mapping programme may be used to assess changes in a singer's voice compared to a healthy baseline or identify specific problem areas across the f_0 and SPL range.

Soft phonation tasks (with lower P_{Sub} and adductory force) can reveal vocal instabilities and early signs of disorder to a further extent than louder utterances (Baker, Purdy, et al., 2023b; D. G. Miller et al., 2001). Bastian et al. (1990) found that auditory-perceptual ratings of quietly sung staccato vocalises and 'Happy Birthday' in head voice were significantly associated with severity of voice condition, and in particular vocal fold swelling. The use of glides or glissandi may also be beneficial to observe vocal behaviours across a range of frequencies in one vocal gesture. When displayed spectrographically, irregularities in vibration at specific points in the singer's range can be located and further investigated. Pre- and post-treatment changes may also

be more readily observed in this manner (Echter-nach et al., 2017).

Robust normative data for singers is currently lacking, particularly for aerodynamic voice measurement such as MPT, PTP, and voicing efficiency ratios. Nonetheless, comparison with a singer's healthy-baseline data (the collection of which should be promoted among clinicians who work with singers) should provide some indication of the significance of any change in repeated measures, particularly when viewed in combination with self-report data.

Although many more tools that could be used for non-invasive clinical assessment of the singing voice have been discussed above, several simple tasks and derived metrics can provide a useful baseline assessment of the singing voice and its function. These are listed below in Table 1, along with their elicitation instructions and theoretical links to vocal function. While many of the tasks can be performed on a comfortable pitch and volume (as is traditionally the case), they should also be performed on a prescribed pitch to mitigate possible f_0 effects. Doing so will improve repeated-measure and intersubject comparability and will also engage the singer in singing rather than speech-like behaviours. The choice of C4 and C5 for lower and higher voices, respectively, is somewhat arbitrary. However, these pitches have been used in previous works and, based on practitioner's personal experience, are not unduly challenging for singers of any voice type or genre. In fact, an *inability to perform tasks* at these frequencies may itself add to their diagnostic utility. The prescribed frequency may change from client to client, depending on their vocal demands or voice type (*Fach*), however this must be noted and reported, and kept consistent between measurement points to ensure comparability.

A *PRAAT script* has been provided in the supplementary material of this work that can be run to extract the acoustic parameters described in Table 1. Although not addressed in the present work, the reader is encouraged to review the following referenced works to guidance in recording setup and SPL calibration (Švec & Granqvist, 2010; Švec & Granqvist, 2018; Winholtz & Titze, 1997). This script is intended to aid clinicians in establishing a baseline of simple measures for singing-voice diagnostics, but these are by no means exhaustive. Other assessments may be useful or needed to determine the nature and extent of a voice complaint and its physiological aetiology. The script provides an initial form into which self-report, aerodynamic, and any other relevant data may be manually added and included in the generated report.

Domain	Measure	Instruction	Interpretation and Notes
Self-Reported Subjective Symptoms	Evaluation of Ability to Sing Easily (EASE)	As per Phyland et al. (2014), reported in Baker et al. (2023).	Changes in Vocal Fatigue, Pathological Risk Indicators, and Voice Concerns subscales may be used to assess singer's perception of their own voice condition, as well as treatment effects. Investigating the responses to specific questions within the scale may provide further insights into the nature of the problem.
	f_0 range (Semi-tones)	'Glide down on an /a/ vowel to your lowest pitch without vocal fry or whispering'; 'Glide up on an /a/ vowel to your highest pitch (head voice included).'	Changes in physiological voice range (i.e., the extent to which the singer can adjust the longitudinal tension of the vocal folds) may be affected by structural change, nerve damage, age-related change, or muscular tensions that inhibit vocal fold lengthening. Singers should be reassured that this measure is not related to their musicianship or voice quality – they should attempt to reach the true minimum and maximum possible frequencies rather than their 'performance' range or 'beautiful' singing.
	Glide through registral transition zone	'Sing a glide/glissando through your registral 'break' from chest-voice to head-voice and back down again.'	Observing the spectrographic data from this task may reveal bifurcation, subharmonics, moments of aphonia, or phonation gaps. While a small 'break' when transitioning from chest to head voice can be common in healthy singers, large breaks with clear timbral changes and instability may suggest a functional or structural voice issues related to the vocal folds' ability to lengthen and thin when moving into head-voice, or conversely, an inability to achieve fuller adduction and greater mass in chest-voice (Behlau et al., 2023; Echternach et al., 2017). The singer's own artistic goals may also inform data interpretation. Performing this task quietly is likely to reveal instabilities to a greater degree than loudly, but the clinician should consider what is appropriate for the singer's style and genre.
Acoustic	SPL range (dB)	'Sing /a/ as softly as you can without whispering on a C4/C5', 'Sing /a/ as loudly as you can without shouting on a C4/C5.'	Changes in intensity range may indicate reduction in ability to regulate P_{sub} , adductory forces, or resonance strategies. Reduced SPL range may result from neurological, structural, or functional voice disorder. In singers, vocal fold oedema from overuse/abuse or substance inhalation may limit SPL range. The task should be sung, that is, not speech-quality phonation, and with vibrato if stylistically relevant. The chosen vowel should be kept consistent, as other acoustic measures (described below) can be affected by resonance-harmonics effects, that is, differences in vowel choice. Here an /a/ is suggested, but whatever vowel is chosen should be kept consistent between measurement points.
	CPPS soft phonation	Taken from 'soft' phonation task above.	Soft phonation may reveal vocal instabilities or inefficiencies to a greater extent than louder phonation. Although CPPS values are affected by SPL and f_0 , by controlling for the f_0 using a prescribed pitch (C4 of C5 for lower or higher voices) and the SPL by obtaining the softest sung tone (without whispering), these effects can be somewhat mitigated. As the SPL of this sample is also given in the report, any significant SPL effects can be taken into account post hoc when required. The prescription of a specific frequency will also ensure comparability across repeated measures as well as between patients.
	L_1 - L_2 soft phonation (C4/C5)	Taken from 'soft' phonation task above.	The level difference between the first two harmonics may provide insights into glottal configuration as well as voicing efficiency related to glottal closure. A more prominent first harmonic (f_1) has been associated with more breathy type phonation, whilst a large negative difference (stronger L_2) suggests phonation toward pressed (Sundberg, 2022; Sundberg et al., 2004).
	Softly sung song (e.g., Happy Birthday)	'Sing Happy Birthday as quietly as possible on the given starting pitch without whispering.'	Evaluation of the spectrogram and long-term average spectrum of a quietly-sung sample will offer insights into vocal stability and condition of the vocal fold mucosa and changes over a short time frame, such as the presence of oedema or pathological vocal fatigue (Bastian et al., 1990).
Acoustic/aerodynamic	Maximum phonation time (C4/C5)	'Inhale maximally and then sing /a/ for as long as you can on a C4/C5 [depending on voice type] until you completely run out of air.'	Maximum phonation time is widely used and can be a useful gauge of a singer's ability to regulate vocal fold adduction and valving. Naturally MPT is affected by the inspired volume, so care should be taken to ensure the singer inspires as close to 100% vital capacity before performing each attempt and expires as much of the volume as possible. In this way, a volume effect may be somewhat mitigated between attempts and measurement points. The task should be performed in triplicate and the best performance recorded.
Aerodynamic	Habitual singing pressures (C4/C5)	'Sing /papapapapa.../ on a C4/C5.'	Assessing the habitual pressures used when singing will help clinicians determine if excessive pressure is contributing to voice complaints. Higher than necessary pressures may cause maladaptive behaviours that increase glottal resistance and recruit inappropriate laryngeal musculature to sustain phonation.
	Phonation threshold pressure (C4/C5)	'Sing /papapapapa.../ on a C4/C5 as quietly as you can without whispering.'	Assessing the minimum pressure required for phonation provides useful data on the condition of the vocal folds, where higher PTP values suggest reduced vocal fold motility, increased mucosal viscosity, and increased vocal fold stiffness, common in oedema. Clients may need multiple trials to achieve true minimum.

Table 1. Possible Voice Measures for Baseline Clinical Singing-Voice Assessment

CONCLUSIONS

Singers are a unique cohort of professional voice users who require a versatile and comprehensive approach to voice analysis. When a singing client presents with a voice complaint, vocal fold visualisation should be the first step. In the absence of laryngeal findings, the use of acoustic and aerodynamic voice analysis can be invaluable for providing additional data on voice production across a range of relevant tasks. These can assist the clinician or voice-rehabilitation specialist in determining problem areas. Although there are a great number of voice analysis techniques available to voice clinicians and researchers, few of these have been systematically tested for use with singing voice samples. Several widely-applied acoustic measures have demonstrated f_0 , SPL, and resonance-harmonics interaction effects that can skew data interpretation, if not accounted for. Furthermore, many tasks used in voice clinics to elicit voice recordings for analysis are aimed at speech-type phonation and do not consider the unique demands for singers.

The use of *soft phonation tasks*, also in combination with glides may show up vocal faults that can be disguised with higher SPLs and firmer vocal fold adduction. *Singing-specific parameters* such as registrational transitions and vibrato can offer invaluable insights into vocal condition during singing and may help guide the clinician in their decision-making when considering treatment approaches. Although not a feature of all singing styles and genres, when vibrato is present, it is generally considered desirable to have a regular and even rate and extent. Deviation from this should be an artistic choice for stylistic or communicative effect rather than a functional default. Noting baseline vibrato characteristics may also help to determine when deviation from such is clinically relevant.

Crucially, the singer's own *perception of their vocal function and condition* must be seriously considered and should be used to guide analysis, as well as when planning treatment goals. *Knowledge of terminology* common in singing voice pedagogy will be useful to interpret singer clients' descriptions of

their voice concerns. Consideration should also be given to the *role of vocal economy* within voice production and when considering treatment options. This should be balanced with indicators of *efficiency* (derived predominantly from the vocal output), and in collaboration with singer reports of *perceived effort* (for example, if the task was more or less difficult than normal). Specifically, the cost of efficiency should be weighed up in light of the whole system including *vocal input parameters* (for example, P_{Sub} , adduction, and collision) and should not be based solely on quantitative acoustic metrics.

More work is needed to establish clinical protocols and normative data that can help identify early biomarkers of at-risk singing voice function or subclinical disorder. In the meantime, existing methods (when properly controlled) such as cepstral, spectral, and spectrographic analysis, as well as 'bigger picture' tools like the VRP or voice mapping should be part of a clinician's regular assessment when analysing a singer's voice. Voice clinicians and medical practitioners should be aware of the unique diagnostic and treatment needs of singers, familiar with the industry-related demands of voice production, and understand the strengths and weaknesses of potential voice analysis approaches. In this way, the functional aetiology of a voice complaint is more likely to be revealed, improving diagnostic accuracy, medical decision-making, and treatment outcomes.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the comments and suggestions of Prof. Dr. Meike Brockmann-Bauser on the initial draft of this work.

FINANCIAL SUPPORT AND SPONSORSHIP

None.

CONFLICT OF INTERESTS

None.

References

- [1] Acosta Martínez, G., & Daffern, H. (2023). Complexity of vocal vibrato in opera and jazz recordings: Insights from entropy and recurrence analyses. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2023.11.020>
- [2] Alku, P., Bäckström, T., & Vilkman, E. (2002). Normalized amplitude quotient for parametrization of the glottal flow. *The Journal of the Acoustical Society of America*, 112(2), 701–710. <https://doi.org/10.1121/1.1490365>
- [3] Allen, J. E. (2019). Editorial: Going with the flow. *Current Opinion in Otolaryngology & Head & Neck Surgery*, 27(6), 431–432. <https://doi.org/10.1097/MOO.0000000000000592>
- [4] Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd ed.). Singular Publishing Group.
- [5] Baker, C. P. (2023). *Female adolescent voice change: Exploring tools for instrumental singing-voice analysis* [University of Auckland]. <https://hdl.handle.net/2292/63774>
- [6] Baker, C. P., Brockmann-Bauser, M., Purdy, S. C., & Rakena, T. O. (2023). High and wide: An in-silico investigation of frequency, intensity, and vibrato effects on widely applied acoustic voice perturbation and noise measures. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2023.10.007>
- [7] Baker, C. P., Miles, A., Allen, J., & Herbst, C. T. (2025). Six-pack singing: On associations between anterior abdominal muscle activity and vital capacity percentage. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2025.03.039>
- [8] Baker, C. P., Purdy, S. C., Rakena, T. O., & Bonnini, S. (2023a). It sounds like it feels: Preliminary exploration of an aeroacoustic diagnostic protocol for singers. *Journal of Clinical Medicine*, 12(15), 5130. <https://doi.org/10.3390/jcm12155130>
- [9] Baker, C. P., Purdy, S. C., Rakena, T. O., & Bonnini, S. (2023b, August 24). An aeroacoustic index for objectively measuring singing-voice fatigue: A proof-of-concept study. *32nd World Congress of the IALP*.
- [10] Baker, C. P., Sundberg, J., Purdy, S. C., Rakena, T. O., & Leão, S. H. de S. (2024). CPPS and voice-source parameters: Objective analysis of the singing voice. *Journal of Voice*, 38(3), 549–560. <https://doi.org/10.1016/j.jvoice.2021.12.010>
- [11] Barlow, C., & LoVetri, J. (2010). Closed quotient and spectral measures of female adolescent singers in different singing styles. *Journal of Voice*, 24(3), 314–318. <https://doi.org/10.1016/j.jvoice.2008.10.003>
- [12] Barsties v Latoszek, B., Maryn, Y., Gerrits, E., & De Bodt, M. (2017). The Acoustic Breathiness Index (ABI): A multivariate acoustic model for breathiness. *Journal of Voice*, 31(4), 511.e11–511.e27. <https://doi.org/10.1016/j.jvoice.2016.11.017>
- [13] Barsties v. Latoszek, B., Müller, A., & Nasr, A. (2025). *Diagnostic and treatment in voice therapy*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-70162-1>
- [14] Bastian, R. W., Keidar, A., & Verdolini-Marston, K. (1990). Simple vocal tasks for detecting vocal fold swelling. *Journal of Voice*, 4(2), 172–183. [https://doi.org/10.1016/S0892-1997\(05\)80144-4](https://doi.org/10.1016/S0892-1997(05)80144-4)
- [15] Behlau, M., Madazio, G., & Yamasaki, R. (2023). Dynamic vocal analysis: Vocal functionality evaluation. *CoDAS*, 35(5). <https://doi.org/10.1590/2317-1782/20232021083en>
- [16] Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* (Version 6.3.10) [Computersoftware]. Phonetic Sciences, University of Amsterdam. <https://praat.org>
- [17] Bonjyotsna, A., & Bhuyan, M. (2016). Analytical study of vocal vibrato and mordent of Indian popular singers. *Journal of Voice*, 30(6), 764.e11–764.e22. <https://doi.org/10.1016/j.jvoice.2015.10.010>
- [18] Bouhuys, A., Mead, J., Proctor, D. F., & Stevens, K. N. (1968). Pressure-flow events during singing. *Annals of the New York Academy of Sciences*, 155(1), 165–176. <https://doi.org/10.1111/j.1749-6632.1968.tb56760.x>
- [19] Bozeman, K. (2013). *Practical vocal acoustics: Pedagogical applications for teachers and singers*. Pendragon Press.
- [20] Brockmann, M., Drinnan, M. J., Storck, C., & Carding, P. N. (2011). Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of Voice*, 25(1), 44–53. <https://doi.org/10.1016/j.jvoice.2009.07.002>
- [21] Brockmann, M., Storck, C., Carding, P. N., & Drinnan, M. J. (2008). Voice loudness and gender effects on jitter and shimmer in healthy adults. *Journal of Speech, Language, and Hearing Research*, 51(5), 1152–1160. [https://doi.org/10.1044/1092-4388\(2008/06-0208\)](https://doi.org/10.1044/1092-4388(2008/06-0208))
- [22] Brockmann-Bauser, M. (2024). Instrumental analysis of voice. In M. J. Ball, N. Müller, & E. Spencer (Eds.), *The Handbook of Clinical Linguistics* (2nd ed., pp. 523–537). Wiley Blackwell. <https://doi.org/10.1002/9781119875949.ch36>
- [23] Brockmann-Bauser, M., Bohlender, J. E., & Mehta, D. D. (2018). Acoustic perturbation measures improve with increasing vocal intensity in individuals with and without voice disorders. *Journal of Voice*, 32(2), 162–168. <https://doi.org/10.1016/j.jvoice.2017.04.008>
- [24] Brockmann-Bauser, M., Van Stan, J. H., Sampaio, M. C., Bohlender, J. E., Hillman, R. E., & Mehta, D. D. (2021). Effects of vocal intensity and fundamental frequency on cepstral peak prominence in patients with voice disorders and vocally healthy controls. *Journal of Voice*, 35(3), 411–417. <https://doi.org/10.1016/j.jvoice.2019.11.015>
- [25] Butte, C. J., Zhang, Y., Song, H., & Jiang, J. J. (2009). Perturbation and nonlinear dynamic analysis of different singing styles. *Journal of Voice*, 23(6), 647–652. <https://doi.org/10.1016/j.jvoice.2008.02.004>
- [26] Calvache Mora, C. A., Cantor-Cutiva, L. C., Hunter, E. J., Guzman, M., & Solaque, L. (2023). Systematic review of literature on vocal demand response: Understanding physiology, measurements, and associated factors. *Folia Phoniatrica et Logopaedica*, 1–1. <https://doi.org/10.1159/000531678>
- [27] Chang, A., & Karnell, M. P. (2004). Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: A study of vocal fatigue. *Journal of Voice*, 18(4), 454–466. <https://doi.org/10.1016/j.jvoice.2004.01.004>
- [28] Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *The Journal of the Acoustical Society of America*, 90(5), 2394–2410. <https://doi.org/10.1121/1.402044>

- [29] Childs, L. F., D'Oto, A., Beams, D. R., Hynan, L., & Mau, T. (2023). Association of genre of singing and phonotraumatic vocal fold lesions in singers. *The Laryngoscope*, 133(7), 1683–1689. <https://doi.org/10.1002/lary.30414>
- [30] Cleveland, T., & Sundberg, J. (1988). Acoustic analysis of three male voices of different quality. *SMAC 83. Proceedings of the Stockholm International Music Acoustics Conference*, 143–156.
- [31] Cohen, S. M., Jacobson, B. H., Garrett, C. G., Noordzij, J. P., Stewart, M. G., Attia, A., Ossoff, R. H., & Cleveland, T. F. (2007). Creation and validation of the singing voice handicap index. *Annals of Otolaryngology, Rhinology & Laryngology*, 116(6), 402–406. <https://doi.org/10.1177/000348940711600602>
- [32] Cohen, S. M., Statham, M., Rosen, C. A., & Zullo, T. (2009). Development and validation of the Singing Voice Handicap-10. *The Laryngoscope*, 119(9), 1864–1869. <https://doi.org/10.1002/lary.20580>
- [33] Dejonckere, P. H., Giordano, A., Schoentgen, J., Fraj, S., Bocchi, L., & Manfredi, C. (2012). To what degree of voice perturbation are jitter measurements valid? A novel approach with synthesized vowels and visuo-perceptual pattern recognition. *Biomedical Signal Processing and Control*, 7(1), 37–42. <https://doi.org/10.1016/j.bspc.2011.05.002>
- [34] DeJonckere, P. H., & Lebacqz, J. (2022). Vocal fold collision speed in vivo: The effect of loudness. *Journal of Voice*, 36(5), 608–621. <https://doi.org/10.1016/j.jvoice.2020.08.025>
- [35] Devin, Z. K. I. (2023). *Vocal fatigue in collegiate singers: An exploratory investigation towards a new assessment protocol* [ProQuest Dissertations & Theses Global]. <https://www.proquest.com/dissertations-theses/vocal-fatigue-collegiate-singers-exploratory/docview/2911044976/se-2>
- [36] D'haeseleer, E., Leyns, C., Meerschman, I., Thyssen, J., Dewaele, F., & Van Lierde, K. (2025). EASE-NL: Cross-cultural adaptation and validation of the Dutch version of the Evaluation of Ability to Sing Easily. *Journal of Voice*, 39(2), 568.e19–568.e27. <https://doi.org/10.1016/j.jvoice.2022.10.003>
- [37] Dippold, S., Voigt, D., Richter, B., & Echternach, M. (2015). High-speed imaging analysis of register transitions in classically and jazz-trained male voices. *Folia Phoniatrica et Logopaedica*, 67(1), 21–28. <https://doi.org/10.1159/000381095>
- [38] Echternach, M., Burk, F., Burdumy, M., Herbst, C. T., Köberlein, M., Döllinger, M., & Richter, B. (2017). The influence of vocal fold mass lesions on the passaggio region of professional singers. *The Laryngoscope*, 127(6), 1392–1401. <https://doi.org/10.1002/lary.26332>
- [39] Enflo, L., Sundberg, J., & McAllister, A. (2013). Collision and phonation threshold pressures before and after loud, prolonged vocalization in trained and untrained voices. *Journal of Voice*, 27(5), 527–530. <https://doi.org/10.1016/j.jvoice.2013.03.008>
- [40] Fraile, R., & Godino-Llorente, J. I. (2014). Cepstral peak prominence: A comprehensive analysis. *Biomed Signal Process Control*, 14, 42–54. <https://doi.org/10.1016/j.bspc.2014.07.001>
- [41] Gauffin, J., & Sundberg, J. (1989). Spectral correlates of glottal voice source waveform characteristics. *Journal of Speech, Language, and Hearing Research*, 32(3), 556–565. <https://doi.org/10.1044/jshr.3203.556>
- [42] Gillivan-Murphy, P., Miller, N., & Carding, P. (2019). Voice tremor in Parkinson's disease: An acoustic study. *Journal of Voice*, 33(4), 526–535. <https://doi.org/10.1016/j.jvoice.2017.12.010>
- [43] Gramming, P., & Sundberg, J. (1988). Spectrum factors relevant to phonetogram measurement. *The Journal of the Acoustical Society of America*, 83(6), 2352–2360. <https://doi.org/10.1121/1.396366>
- [44] Granqvist, S. (2025a). *Description and validation of RecVox – A free software for real-time phonetography*. https://www.tolvan.com/recvox/Validation_of_RecVox.pdf
- [45] Granqvist, S. (2025b). *RecVox*. Tolvan Data. <https://www.tolvan.com/index.php?page=/recvox/recvox.php>
- [46] Granqvist, S. (2025c). *Sopran*. Tolvan Data. www.tolvan.com
- [47] Gunjawate, D. R., Ravi, R., & Bellur, R. (2018). Acoustic analysis of voice in singers: A systematic review. *Journal of Speech, Language, and Hearing Research*, 61, 40–51. https://doi.org/10.1044/2017_JSLHR-S-17-0145
- [48] Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J., & Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Oto-Laryngologica*, 90(1–6), 441–451. <https://doi.org/10.3109/00016488009131746>
- [49] Hassan, E. M., Abdel Hady, A. F., Shohdi, S. S., Eldessouky, H. M., & Din, M. H. B. (2021). Assessment of dysphonia: Cepstral analysis versus conventional acoustic analysis. *Logopedics Phoniatrics Vocology*, 46(3), 99–109. <https://doi.org/10.1080/14015439.2020.1767202>
- [50] Hejná, M., Šturm, P., Tylečková, L., & Bořil, T. (2019). Normo-phonetic breathiness in Czech and Danish: Are females breathier than males? *Journal of Voice*, 35(3), 498.e1–498.e22. <https://doi.org/10.1016/j.jvoice.2019.10.019>
- [51] Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Baroody, M. M., Hillenbrand, J., & Sataloff, R. T. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otolaryngology, Rhinology & Laryngology*, 112(4), 324–333. <https://doi.org/10.1177/000348940311200406>
- [52] Herbst, C. T. (2017). A review of singing voice subsystem interactions—Toward an extended physiological model of “support.” *Journal of Voice*, 31(2), 249.e13–249.e19. <https://doi.org/10.1016/j.jvoice.2016.07.019>
- [53] Herbst, C. T., Hess, M., Müller, F., Švec, J. G., & Sundberg, J. (2015). Glottal adduction and subglottal pressure in singing. *Journal of Voice*, 29(4), 391–402. <https://doi.org/10.1016/j.jvoice.2014.08.009>
- [54] Herbst, C. T., Unger, J., Herzel, H., Švec, J. G., & Lohscheller, J. (2016). Phasegram analysis of vocal fold vibration documented with laryngeal high-speed video endoscopy. *Journal of Voice*, 30(6), 771.e1–771.e15. <https://doi.org/10.1016/j.jvoice.2015.11.006>
- [55] Hertegård, S., Gauffin, J., & Lindestad, P.-Å. (1995). A comparison of subglottal and intraoral pressure measurements during phonation. *Journal of Voice*, 9(2), 149–155. [https://doi.org/10.1016/S0892-1997\(05\)80248-6](https://doi.org/10.1016/S0892-1997(05)80248-6)
- [56] Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech,*

- Language, and Hearing Research*, 37(4), 769–779. <https://doi.org/10.1044/jshr.3704.769>
- [57] Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research*, 39(2), 311–321. <https://doi.org/10.1044/jshr.3902.311>
- [58] Hillman, R. E., Stepp, C. E., Van Stan, J. H., Zañartu, M., & Mehta, D. D. (2020). An updated theoretical framework for vocal hyperfunction. *American Journal of Speech-Language Pathology*, 29(4), 2254–2260. https://doi.org/10.1044/2020_AJSLP-20-00104
- [59] Horii, Y. (1989a). Acoustic analysis of vocal vibrato: A theoretical interpretation of data. *Journal of Voice*, 3(1), 36–43. [https://doi.org/10.1016/S0892-1997\(89\)80120-1](https://doi.org/10.1016/S0892-1997(89)80120-1)
- [60] Horii, Y. (1989b). Frequency modulation characteristics of sustained /a/ sung in vocal vibrato. *Journal of Speech, Language, and Hearing Research*, 32(4), 829–836. <https://doi.org/10.1044/jshr.3204.829>
- [61] Hunter, E. J., Cantor-Cutiva, L. C., van Leer, E., van Mersbergen, M., Nanjundeswaran, C. D., Bottalico, P., Sandage, M. J., & Whitling, S. (2020). Toward a consensus description of vocal effort, vocal load, vocal loading, and vocal fatigue. *Journal of Speech, Language, and Hearing Research*, 63(2), 509–532. https://doi.org/10.1044/2019_JSLHR-19-00057
- [62] Iwata, S., Von Leden, H., & Williams, D. (1972). Air flow measurement during phonation. *Journal of Communication Disorders*, 5(1), 67–79. [https://doi.org/10.1016/0021-9924\(72\)90033-0](https://doi.org/10.1016/0021-9924(72)90033-0)
- [63] KayPentax. (n.d.-a). *Multi-Dimensional Voice Program (MDVP)*. Kay Elemetrics Corporation.
- [64] KayPentax. (n.d.-b). *Phonatory Aerodynamic System (PAS)*. Kay Elemetrics Corporation.
- [65] Kob, M., Henrich, N., Herzel, H., Howard, D., Tokuda, I., & Wolfe, J. (2011). Analysing and understanding the singing voice: Recent progress and open questions. *Current Bioinformatics*, 6(3), 362–374. <https://doi.org/10.2174/157489311796904709>
- [66] Lã, F. M. B., & Fiuza, M. B. (2022). Real-time visual feedback in singing pedagogy: Current trends and future directions. *Applied Sciences*, 12(21), 10781. <https://doi.org/10.3390/app122110781>
- [67] Lã, F. M. B., Silva, L. S., & Granqvist, S. (2023). Long-Term Average Spectrum Characteristics of Portuguese Fado-Canção from Coimbra. *Journal of Voice*, 37(4), 631.e7–631.e15. <https://doi.org/10.1016/j.jvoice.2021.03.005>
- [68] Ladefoged, P., & McKinney, N. P. (1963). Loudness, sound pressure, and subglottal pressure in speech. *The Journal of the Acoustical Society of America*, 35(4), 454–460. <https://doi.org/10.1121/1.1918503>
- [69] Lamarche, A. (2009). *Putting the singing voice on the map* [Doctoral Thesis, KTH]. https://www.speech.kth.se/~anick/elec_tmh.pdf
- [70] Lamarche, A., Ternström, S., & Hertegård, S. (2009). Not just sound: Supplementing the voice range profile with the singer's own perceptions of vocal challenges. *Logopedics Phoniatrics Vocology*, 34(1), 3–10. <https://doi.org/10.1080/14015430802239759>
- [71] Lamarche, A., Ternström, S., & Pabon, P. (2010). The singer's voice range profile: Female professional opera soloists. *Journal of Voice*, 24(4), 410–426. <https://doi.org/10.1016/j.jvoice.2008.12.008>
- [72] Lechien, J. R., Geneid, A., Bohlender, J. E., Cantarella, G., Avellaneda, J. C., Desuter, G., Sjogren, E. V., Finck, C., Hans, S., Hess, M., Oguz, H., Remacle, M. J., Schneider-Stickler, B., Tedla, M., Schindler, A., Vilaseca, I., Zabrodsky, M., Dikkers, F. G., & Crevier-Buchman, L. (2023). Consensus for voice quality assessment in clinical practice: guidelines of the European Laryngological Society and Union of the European Phoniatricians. *European Archives of Oto-Rhino-Laryngology*. <https://doi.org/10.1007/s00405-023-08211-6>
- [73] Lycke, H., & Siupsinskiene, N. (2016). Voice range profiles of singing students: The effects of training duration and institution. *Folia Phoniatrica et Logopaedica*, 68(2), 53–59. <https://doi.org/10.1159/000448136>
- [74] Manfredi, C., Barbagallo, D., Baracca, G., Orlandi, S., Bandini, A., & Dejonckere, P. H. (2015). Automatic assessment of acoustic parameters of the singing voice: Application to professional western operatic and jazz singers. *Journal of Voice*, 29(4), 517.e1–517.e9. <https://doi.org/10.1016/j.jvoice.2014.09.014>
- [75] Manfredi, C., Giordano, A., Schoentgen, J., Fraj, S., Bocchi, L., & Dejonckere, P. H. (2012). Perturbation measurements in highly irregular voice signals: Performances/validity of analysis software tools. *Biomedical Signal Processing and Control*, 7(4), 409–416. <https://doi.org/10.1016/j.bspc.2011.06.004>
- [76] Maryn, Y., De Bodt, M., Barsties, B., & Roy, N. (2013). The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology*. <https://doi.org/10.1007/s00405-013-2730-7>
- [77] Maryn, Y., De Bodt, M., & Roy, N. (2010). The Acoustic Voice Quality Index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, 43(3), 161–174. <https://doi.org/10.1016/j.jcomdis.2009.12.004>
- [78] Maryn, Y., Dwenger, K., Kaufmann, S., & Barkmeier-Kraemer, J. (2025). Reliability and diagnostic accuracy of semi-automated and automated acoustic quantification of vocal tremor characteristics. *Journal of Speech, Language, and Hearing Research*, 1–20. https://doi.org/10.1044/2025_JSLHR-24-00467
- [79] Maryn, Y., Leblans, M., Zarowski, A., & Barkmeier-Kraemer, J. (2019). Objective acoustic quantification of perceived voice tremor severity. *Journal of Speech, Language, and Hearing Research*, 62(10), 3689–3705. https://doi.org/10.1044/2019_JSLHR-S-19-0024
- [80] Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program PRAAT: Smoothed cepstral peak prominence and Acoustic Voice Quality Index. *Journal of Voice*, 29(1), 35–43. <https://doi.org/10.1016/j.jvoice.2014.06.015>
- [81] McCoy, S. J. (2012). *Your Voice: An inside view* (2nd ed.). Inside View Press.
- [82] Miller, D. G., Schutte, H. K., & Doing, J. (2001). Soft phonation in the male singing voice. *Journal of Voice*, 15(4), 483–491. [https://doi.org/10.1016/S0892-1997\(01\)00048-0](https://doi.org/10.1016/S0892-1997(01)00048-0)

- [83] Miller, R. (1996). *The structure of singing: System and art in vocal technique*. Schirmer.
- [84] Mitchell, H. F., & Kenny, D. T. (2010). Change in vibrato rate and extent during tertiary training in classical singing students. *Journal of Voice*, 24(4), 427–434. <https://doi.org/10.1016/j.jvoice.2008.12.003>
- [85] Morelli, M. S., Orlandi, S., & Manfredi, C. (2021). BioVoice: A multipurpose tool for voice analysis. *Biomedical Signal Processing and Control*, 64, 102302. <https://doi.org/10.1016/j.bspc.2020.102302>
- [86] Mori, M. C., Francis, D. O., & Song, P. C. (2017). Identifying occupations at risk for laryngeal disorders requiring specialty voice care. *Otolaryngology–Head and Neck Surgery*, 157(4), 670–675. <https://doi.org/10.1177/0194599817726528>
- [87] Murray, E. H., & Woodnorth, G. H. (2020). Clinical approach to acoustic assessment. In J. S. McMurray, M. R. Hoffmann, & M. N. Braden (Eds.), *Multidisciplinary management of pediatric voice and swallowing disorders* (pp. 83–88). Springer. https://doi.org/10.1007/978-3-030-26191-7_9
- [88] Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *American Journal of Speech-Language Pathology*, 29(3), 1596–1607. https://doi.org/10.1044/2020_AJSLP-20-00001
- [89] Nandamudi, S. (2017). *Aerodynamics of Vocal Vibrato* [Doctoral dissertation, Bowling Green State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=bgsu1499427478103556
- [90] Nestorova, T., Brandner, M., Gingras, B., & Herbst, C. T. (2023). Vocal vibrato characteristics in historical and contemporary opera, operetta, and Schlager. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2022.12.027>
- [91] Nix, J., Perna, N., James, K., & Allen, S. (2016). Vibrato rate and extent in college music majors: A multicenter study. *Journal of Voice*, 30(6), 756.e31–756.e41. <https://doi.org/10.1016/j.jvoice.2015.09.006>
- [92] Nordenberg, M., & Sundberg, J. (2004). Effect on LTAS of vocal loudness variation. *Logopedics Phoniatrics Vocology*, 29(4), 183–191. <https://doi.org/10.1080/14015430410004689>
- [93] Oates, J. M., Bain, B., Davis, P., Chapman, J., & Kenny, D. (2006). Development of an auditory-perceptual rating instrument for the operatic singing voice. *Journal of Voice*, 20(1), 71–81. <https://doi.org/10.1016/j.jvoice.2005.01.006>
- [94] Öğülmüş Uysal, M., Esen Aydinli, F., İncebay, Ö., Beşik Topçu, Ö., & Daşdoğan, Ü. (2024). Cross-cultural adaptation and validation of the Turkish version of the Evaluation of the Ability to Sing Easily (EASE-TR). *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2024.10.016>
- [95] Otori, K., Kojima, H., Kakani, R., Slavik, D. H., & Blaugrund, S. M. (1997). Acoustic characteristics of rough voice: Subharmonics. *Journal of Voice*, 11(1), 40–47. [https://doi.org/10.1016/S0892-1997\(97\)80022-7](https://doi.org/10.1016/S0892-1997(97)80022-7)
- [96] Orlandi, S., Dejonckere, P. H., Schoentgen, J., Lebacqz, J., Rruqja, N., & Manfredi, C. (2013). Effective pre-processing of long term noisy audio recordings: An aid to clinical monitoring. *Biomedical Signal Processing and Control*, 8(6), 799–810. <https://doi.org/10.1016/j.bspc.2013.07.009>
- [97] Orlikoff, R. F., & Kahane, J. C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. *Journal of Voice*, 5(2), 113–119. [https://doi.org/10.1016/S0892-1997\(05\)80175-4](https://doi.org/10.1016/S0892-1997(05)80175-4)
- [98] Pabon, P. (1991). Objective acoustic voice-quality parameters in the computer phonetogram. *Journal of Voice*, 5(3), 203–216. [https://doi.org/10.1016/S0892-1997\(05\)80188-2](https://doi.org/10.1016/S0892-1997(05)80188-2)
- [99] Pabon, P., Stallinga, R., Södersten, M., & Ternström, S. (2014). Effects on vocal range and voice quality of singing voice training: The classically trained female voice. *Journal of Voice*, 28(1), 36–51. <https://doi.org/10.1016/j.jvoice.2013.06.005>
- [100] Pabon, P., & Ternström, S. (2020). Feature maps of the acoustic spectrum of the voice. *Journal of Voice*, 34(1), 161.e1–161.e26. <https://doi.org/10.1016/j.jvoice.2018.08.014>
- [101] Pabon, P., Ternström, S., & Lamarche, A. (2011). Fourier descriptor analysis and unification of voice range profile contours: Method and applications. *Journal of Speech, Language, and Hearing Research*, 54(3), 755–776. [https://doi.org/10.1044/1092-4388\(2010/08-0222\)](https://doi.org/10.1044/1092-4388(2010/08-0222))
- [102] Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Coury, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology*, 27, 887–905. https://doi.org/10.1044/2018_AJSLP-17-0009
- [103] Patel, R. R., Sandage, M. J., & Golzarri-Arroyo, L. (2023). High-speed videoendoscopic and acoustic characteristics of inspiratory phonation. *Journal of Speech, Language, and Hearing Research*, 66(4), 1192–1207. https://doi.org/10.1044/2022_JSLHR-22-00502
- [104] Perrin, C. E., Young, V. N., Ma, Y., Rosen, C. A., Stockton, S. D., & Schneider, S. L. (2025). Singing Voice Handicap Index-10 minimal clinically important difference: A prospective determination. *The Laryngoscope*, 135(2), 752–757. <https://doi.org/10.1002/lary.31808>
- [105] Pestana, P. M., Vaz-Freitas, S., & Manso, M. C. (2017). Prevalence of voice disorders in singers: Systematic review and meta-analysis. *Journal of Voice*, 31(6), 722–727. <https://doi.org/10.1016/j.jvoice.2017.02.010>
- [106] Pestana, P. M., Vaz-Freitas, S., & Manso, M. C. (2023). Risk factors for voice disorders among Fado singers: A cross-sectional study. *Journal of Otorhinolaryngology, Hearing and Balance Medicine*, 4(1), 5. <https://doi.org/10.3390/ohbm4010005>
- [107] Phyland, D. J. (2014). *The impact of vocal load on the vocal function of professional music theatre singers*. [Doctoral Dissertation, Monash University] <https://doi.org/10.4225/03/58b77627ae172>
- [108] Phyland, D. J. (2017). The measurement and effects of vocal load in singing performance. How much singing can a singer sing if a singer can sing songs? *Perspectives of the ASHA Special Interest Groups*, 2(3), 79–88. <https://doi.org/10.1044/persp2.SIG3.79>
- [109] Phyland, D. J., & Miles, A. (2019). Occupational voice is a work in progress: Active risk management, habilitation and rehabilitation. *Current Opinion in Otolaryngology & Head & Neck Surgery*, 27(6), 439–447. <https://doi.org/10.1097/MOO.0000000000000584>

- [110] Phyland, D. J., Pallant, J. F., Benninger, M. S., Thibeault, S. L., Greenwood, K. M., Smith, J. A., & Vallance, N. (2013). Development and preliminary validation of the EASE: A tool to measure perceived singing voice function. *Journal of Voice*, 27(4), 454–462. <https://doi.org/10.1016/j.jvoice.2013.01.019>
- [111] Phyland, D. J., Pallant, J. F., Thibeault, S. L., Benninger, M. S., Vallance, N., & Smith, J. A. (2014). Measuring vocal function in professional music theater singers: Construct validation of the Evaluation of the Ability to Sing Easily (EASE). *Folia Phoniatrica et Logopaedica*, 66(3), 100–108. <https://doi.org/10.1159/000366202>
- [112] Plexico, L. W., Sandage, M. J., & Faver, K. Y. (2011). Assessment of phonation threshold pressure: A critical review and clinical implications. *American Journal of Speech-Language Pathology*, 20(4), 348–366. [https://doi.org/10.1044/1058-0360\(2011/10-0066\)](https://doi.org/10.1044/1058-0360(2011/10-0066))
- [113] Powell, M. E., Deliyski, D. D., Zeitels, S. M., Burns, J. A., Hillman, R. E., Gerlach, T. T., & Mehta, D. D. (2020). Efficacy of videostroboscopy and high-speed videoendoscopy to obtain functional outcomes from perioperative ratings in patients with vocal fold mass lesions. *Journal of Voice*, 34(5), 769–782. <https://doi.org/10.1016/j.jvoice.2019.03.012>
- [114] Prebil, N., Hočevár Boltežar, I., & Šereg Bahar, M. (2020). Risk factors for voice problems in professional actors and singers. *Slovenian Journal of Public Health*, 59(2), 92–98. <https://doi.org/10.2478/sjph-2020-0012>
- [115] Ramig, L. A., & Shipp, T. (1987). Comparative measures of vocal tremor and vocal vibrato. *Journal of Voice*, 1(2), 162–167. [https://doi.org/10.1016/S0892-1997\(87\)80040-1](https://doi.org/10.1016/S0892-1997(87)80040-1)
- [116] Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America*, 53(6), 1632–1645. <https://doi.org/10.1121/1.1913513>
- [117] Rruqja, N., Dejonckere, P. H., Cantarella, G., Schoentgen, J., Orlandi, S., Barbagallo, S. D., & Manfredi, C. (2014). Testing software tools with synthesized deviant voices for medicolegal assessment of occupational dysphonia. *Biomedical Signal Processing and Control*, 13, 71–78. <https://doi.org/10.1016/j.bspc.2014.03.011>
- [118] Sampaio, M. C., Bohlender, J. E., & Brockmann-Bauser, M. (2021). Fundamental frequency and intensity effects on cepstral measures in vowels from connected speech of speakers with voice disorders. *Journal of Voice*, 35(3), 422–431. <https://doi.org/10.1016/j.jvoice.2019.11.014>
- [119] Sampaio, M. C., Masson, M. L. V., Soares, M. F. de P., Bohlender, J. E., & Brockmann-Bauser, M. (2020). Effects of fundamental frequency, vocal intensity, sample duration, and vowel context in cepstral and spectral measures of dysphonic voices. *Journal of Speech, Language, and Hearing Research*, 63(5), 1326–1339. https://doi.org/10.1044/2020_JSLHR-19-00049
- [120] Sangiorgi, T., Manfredi, C., & Brusciagioni, P. (2009). Objective analysis of the singing voice as a training aid. *Logopedics Phoniatrics Vocology*, 30(3–4), 136–146. <https://doi.org/10.1080/14015430500294064>
- [121] Sataloff, R. T. (1981). Professional singers: The science and art of clinical care. *American Journal of Otolaryngology*, 2(3), 251–266. [https://doi.org/10.1016/S0196-0709\(81\)80022-1](https://doi.org/10.1016/S0196-0709(81)80022-1)
- [122] Sataloff, R. T. (2017). *Clinical assessment of voice* (2nd ed.). Plural Publishing Inc.
- [123] Sataloff, R. T., Hawkshaw, M. J., Johnson, J. L., Ruel, B., Wilhelm, A., & Lurie, D. (2012). Prevalence of abnormal laryngeal findings in healthy singing teachers. *Journal of Voice*, 26(5), 577–583. <https://doi.org/10.1016/j.jvoice.2012.01.002>
- [124] Schultz, B. G., & Vogel, A. P. (2022). A tutorial review on clinical acoustic markers in speech science. *Journal of Speech, Language, and Hearing Research*, 65(9), 3239–3263. https://doi.org/10.1044/2022_JSLHR-21-00647
- [125] Scotto Di Carlo, N. (1994). Internal voice sensitivities in opera singers. *Folia Phoniatrica et Logopaedica*, 46(2), 79–85. <https://doi.org/10.1159/000266296>
- [126] Shembel, A. C., & Nanjundeswaran, C. (2022). Potential biophysiological mechanisms underlying vocal demands and vocal fatigue. *Journal of Voice*, (Advance online publication). <https://doi.org/10.1016/j.jvoice.2022.07.017>
- [127] Shipp, T., Doherty, E. T., & Haglund, S. (1990). Physiologic factors in vocal vibrato production. *Journal of Voice*, 4(4), 300–304. [https://doi.org/10.1016/S0892-1997\(05\)80045-1](https://doi.org/10.1016/S0892-1997(05)80045-1)
- [128] Smitheran, J. R., & Hixon, T. J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders*, 46(2), 138–146. <https://doi.org/10.1044/jshd.4602.138>
- [129] Stevens, K. N. (2000). *Acoustic phonetics*. MIT Press.
- [130] Sundberg, J. (1974). Articulatory interpretation of the “singing formant.” *Journal of the Acoustical Society of America*, 55(4), 838–844. <https://doi.org/10.1121/1.1914609>
- [131] Sundberg, J. (1987). *The science of the singing voice*. Northern Illinois University Press.
- [132] Sundberg, J. (1995). Acoustic and psychoacoustic aspects of vocal vibrato. In P. Dejonckere, M. Hirano, & J. Sundberg (Eds.), *Vibrato* (pp. 35–62). Singular Publishing Company.
- [133] Sundberg, J. (2001). Level and center frequency of the singer’s formant. *Journal of Voice*, 15(2), 176–186. [https://doi.org/10.1016/S0892-1997\(01\)00019-4](https://doi.org/10.1016/S0892-1997(01)00019-4)
- [134] Sundberg, J. (2022). Objective characterization of phonation type using amplitude of flow glottogram pulse and of voice source fundamental. *Journal of Voice*, 36(1), 4–14. <https://doi.org/10.1016/j.jvoice.2020.03.018>
- [135] Sundberg, J., Andersson, M., & Hultqvist, C. (1999). Effects of subglottal pressure variation on professional baritone singers’ voice sources. *The Journal of the Acoustical Society of America*, 105(3), 1965–1971. <https://doi.org/10.1121/1.426731>
- [136] Sundberg, J., Fahlstedt, E., & Morell, A. (2005). Effects on the glottal voice source of vocal loudness variation in untrained female and male voices. *The Journal of the Acoustical Society of America*, 117(2), 879–885. <https://doi.org/10.1121/1.1841612>
- [137] Sundberg, J., Gramming, P., & LoVetri, J. (1993). Comparisons of pharynx, source, formant, and pressure characteristics in operatic and musical theatre singing. *Journal of Voice*, 7(4), 301–310. [https://doi.org/10.1016/S0892-1997\(05\)80118-3](https://doi.org/10.1016/S0892-1997(05)80118-3)
- [138] Sundberg, J., Lã, F. M. B., & Gill, B. P. (2013). Formant tuning strategies in professional male opera singers. *Journal of Voice*, 27(3), 278–288. <https://doi.org/10.1016/j.jvoice.2012.12.002>

- [139] Sundberg, J., & Nordenberg, M. (2006). Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech. *The Journal of the Acoustical Society of America*, 120(1), 453–457. <https://doi.org/10.1121/1.2208451>
- [140] Sundberg, J., Thalén, M., Alku, P., & Vilkman, E. (2004). Estimating perceived phonatory pressedness in singing from flow glottograms. *Journal of Voice*, 18(1), 56–62. <https://doi.org/10.1016/j.jvoice.2003.05.006>
- [141] Švec, J. G., & Granqvist, S. (2010). Guidelines for selecting microphones for human voice production research. *American Journal of Speech-Language Pathology*, 19(4), 356–368. [https://doi.org/10.1044/1058-0360\(2010/09-0091\)](https://doi.org/10.1044/1058-0360(2010/09-0091))
- [142] Švec, J. G., & Granqvist, S. (2018). Tutorial and guidelines on measurement of sound pressure level in voice and speech. *Journal of Speech, Language, and Hearing Research*, 61(3), 441–461. https://doi.org/10.1044/2017_JSLHR-S-17-0095
- [143] Ternström, S. (2005). Does the acoustic waveform mirror the voice? *Logopedics Phoniatrics Vocology*, 30(3–4), 100–107. <https://doi.org/10.1080/14015430500238400>
- [144] Ternström, S., Johansson, D., & Selamtzis, A. (2018). FonaDyn – A System for real-time analysis of the electroglottogram, over the voice range. *SoftwareX*, 7, 74–80. <https://doi.org/10.1016/j.softx.2018.03.002>
- [145] Ternström, S., & Pabon, P. (2022). Voice maps as a tool for understanding and dealing with variability in the voice. *Applied Sciences*, 12(22), 11353. <https://doi.org/10.3390/app122211353>
- [146] Thalén, M., & Sundberg, J. (2001). Describing different styles of singing: A comparison of a female singer's voice source in “Classical”, “Pop”, “Jazz” and “Blues”. *Logopedics Phoniatrics Vocology*, 26(2), 82–93. <https://doi.org/10.1080/140154301753207458>
- [147] Titze, I. R. (1992). Phonation threshold pressure: A missing link in glottal aerodynamics. *The Journal of the Acoustical Society of America*, 91(5), 2926–2935. <https://doi.org/10.1121/1.402928>
- [148] Titze, I. R. (1995). *Workshop on acoustic voice analysis: Summary statement* (N. C. for V. and Speech, Ed.). National Center for Voice and Speech.
- [149] Titze, I. R. (2000). *Principles of voice production*. National Center for Voice and Speech.
- [150] Titze, I. R., Horii, Y., & Scherer, R. C. (1987). Some technical considerations in voice perturbation measurements. *Journal of Speech and Hearing Research*, 30, 252–260. <https://doi.org/10.1044/jshr.3002.252>
- [151] Titze, I. R., & Liang, H. (1993). Comparison of f_0 extraction methods for high-precision voice perturbation measures. *Journal of Speech, Language, and Hearing Research*, 36(6), 1120–1133. <https://doi.org/10.1044/jshr.3606.1120>
- [152] Titze, I. R., Schmidt, S. S., & Titze, M. R. (1995). Phonation threshold pressure in a physical model of the vocal fold mucosa. *The Journal of the Acoustical Society of America*, 97(5), 3080–3084. <https://doi.org/10.1121/1.411870>
- [153] Titze, I. R., Story, B., Smith, M., & Long, R. (2002). A reflex resonance model of vocal vibrato. *The Journal of the Acoustical Society of America*, 111(5), 2272. <https://doi.org/10.1121/1.1434945>
- [154] Titze, I. R., & Sundberg, J. (1992). Vocal intensity in speakers and singers. *The Journal of the Acoustical Society of America*, 91(5), 2936–2946. <https://doi.org/10.1121/1.402929>
- [155] Toles, L. E., Seidman, A. Y., Hillman, R. E., & Mehta, D. D. (2022). Clinical utility of the ratio of sound pressure level to subglottal pressure in patients surgically treated for phonotraumatic vocal fold lesions. *Journal of Speech, Language, and Hearing Research*, 65(8), 2778–2788. https://doi.org/10.1044/2022_JSLHR-21-00658
- [156] Van Houtte, E., Van Lierde, K., D’Haeseleer, E., & Claeys, S. (2010). The prevalence of laryngeal pathology in a treatment-seeking population with dysphonia. *The Laryngoscope*, 120(2), 306–312. <https://doi.org/10.1002/lary.20696>
- [157] Weiss, R., Brown, W. S., & Moris, J. (2001). Singer's formant in sopranos: Fact or fiction? *Journal of Voice*, 15(4), 457–468. [https://doi.org/10.1016/S0892-1997\(01\)00046-7](https://doi.org/10.1016/S0892-1997(01)00046-7)
- [158] Welham, N. v. (2009). Clinical voice evaluation. In A. E. Aronson & D. M. Bless (Eds.), *Clinical voice disorders* (4th ed.). Theime Medical Publishers, Inc.
- [159] Winholtz, W. S., & Titze, I. R. (1997). Conversion of a head-mounted microphone signal into calibrated SPL units. *Journal of Voice*, 11(4), 417–421. [https://doi.org/10.1016/S0892-1997\(97\)80037-9](https://doi.org/10.1016/S0892-1997(97)80037-9)
- [160] Wuyts, F. L., De Bodt, M., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Lierde, K. Van, Raes, J., & Heyning, P. H. Van de. (2000). The Dysphonia Severity Index. *Journal of Speech, Language, and Hearing Research*, 43(3), 796–809. <https://doi.org/10.1044/jshr.4303.796>
- [161] Yamauchi, A., Yokonishi, H., Imagawa, H., Sakakibara, K.-I., Nito, T., Tayama, N., & Yamasoba, T. (2016). Quantification of vocal fold vibration in various laryngeal disorders using high-speed digital imaging. *Journal of Voice*, 30(2), 205–214. <https://doi.org/10.1016/j.jvoice.2015.04.016>

SUPPLEMENTARY MATERIALS

Data collection and extraction for acoustic singing-voice assessment script

Acoustic voice samples should be recorded in a quiet room or booth, ideally with an omnidirectional headset condenser microphone to avoid changes in mic-to-mouth distance. A calibration task should be performed first, where the client sustains and /a/ vowel at a comfortable pitch and volume as steadily as possible, while the clinician records the SPL at 30cm in front of the participant. The level difference between the PRAAT measured intensity (Get intensity...) and the SPL value can then be used as a calibration factor, entered into the initial script form. All SPL values provided in the final report will then reflect SPL @ 30cm. If no

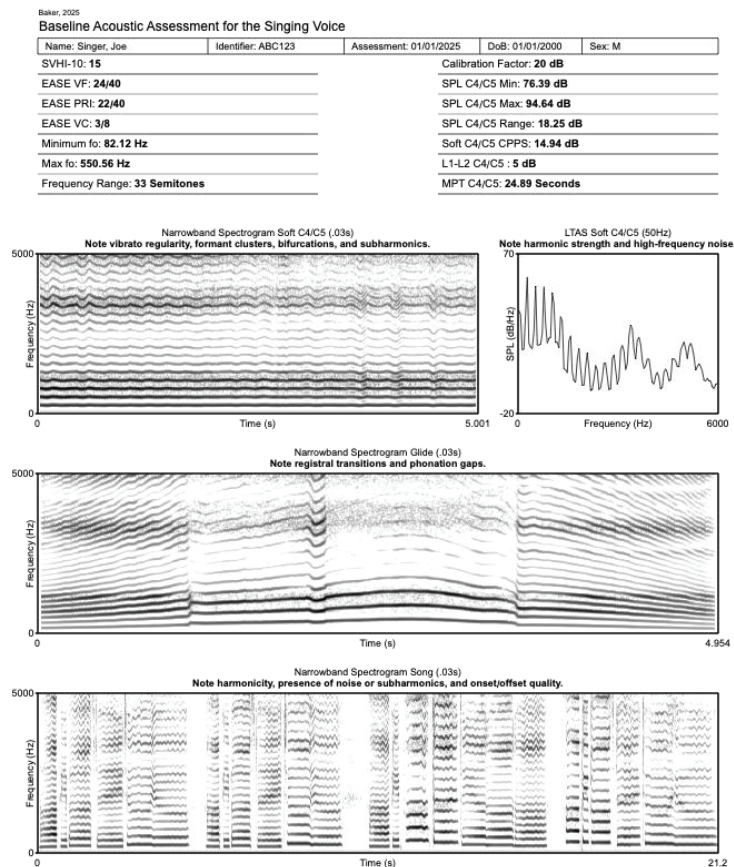


Figure 5. Sample Report Output

calibration is performed, a value of 0 should be entered into the calibration factor box; note that uncalibrated SPL data cannot be compared between measurement points or with other participants. For a thorough discussion on microphone selection for voice analysis and calibration procedures, see Švec and Granqvist (2010; 2018). Sample recordings for testing the script provided below can be downloaded from <https://doi.org/10.17608/k6.auckland.29994499>

Extraction of individual samples (tasks as per Table 1) should be made at zero crossings, and named as follows (without capitalisation or spaces):

- minfo – portion of lowest f_o (do not include vocal fry or silences)
- maxfo – portion of highest f_o (do not include silences, and ensure pitch settings are high enough to accommodate highest pitch (for example, 1500Hz in high-voiced singers)
- glide – full sample or relevant portion of a glide from low to high and back, through a registral transition
- soft – most stable inner portion of softest sung phonation on C4/C5, avoid onset and offset artefacts
- loud – most stable inner portion of loudest sung phonation on C4/C5, avoid onset and offset artefacts
- song – entire sample of a softly-sung song
- mpt – entire maximum phonation time sample from phonation onset to phonation offset

Script

form Acoustic Assessment of the Singing Voice

comment This is a script to calculate acoustic measures for clinical assessment

comment of the singing voice (Baker, 2025). Tasks include minimum and maximum pitch,

comment glide through registral transition, minimum and maximum intensity (C4/C5),

comment maximum phonation time (C4/C5), and a singing sample.

comment

comment See Baker (2025; DOI: 10.2478/ijvs-2025-0003) for task elicitation details. Individual file names

comment should be: calibration, minfo, maxfo, glide, soft, loud, mpt, and song,

comment all without spaces or capitalisation.

comment

comment The dB level difference between the PRAAT measured intensity of a calibration

comment tone and the SPL-meter measurement should be entered into the 'Calibration factor'

comment field. If no calibration was made, leave as 0. For more information on SPL

comment calibration, see Švec and Granqvist (2018; DOI: 10.1044/2017_JSLHR-S-17-0095)

comment

comment Please enter the following information

sentence Assessment_date dd/mm/yyyy

sentence Name

sentence Identifier iD

sentence Sex_if_relevant NA

sentence Date_of_Birth dd/mm/yyyy

real SVHI_10 0

real EASE_vocal_fatigue_score 1

real EASE_pathological_risk_indicators_score 1

real EASE_voice_concerns_score 1

real Calibration_factor 0

text Include_perceptual_features_or_data_from_other_measures_below endform

easevf = eASE_vocal_fatigue_score

easepri = eASE_pathological_risk_indicators_score

easevc = eASE_voice_concerns_score

calfac = calibration_factor

calfacFix\$ = fixed\$ (calfac, 3)

calfacFix = number (calfacFix\$)

svhi10 = sVHI_10

Erase all

Select inner viewport: 0.5, 7.5, 0.5, 4.5

Axes: 0, 1, 0, 1

Colour: "{0,0,0}"

Text special: 0.5, "centre", 0.6, "half", "Palatino", 12, "0", "The data are being calculated. Please wait."

#Analyses

#Multiply to reflect SPL@30cm

selectObject: "Sound minfo"

Copy: "minfoc"
Multiply... $10^{(\text{calfacFix}/20)}$

select Sound maxfo
Copy: "maxfoc"
Multiply... $10^{(\text{calfacFix}/20)}$

select Sound soft
Copy: "softc"
Multiply... $10^{(\text{calfacFix}/20)}$

select Sound loud
Copy: "loudc"
Multiply... $10^{(\text{calfacFix}/20)}$

select Sound glide
Copy: "glidec"
Multiply... $10^{(\text{calfacFix}/20)}$

select Sound song
Copy: "songc"
Multiply... $10^{(\text{calfacFix}/20)}$

select Sound mpt
Copy: "mptc"
Multiply... $10^{(\text{calfacFix}/20)}$

#fo Range

select Sound minfo
To Pitch (raw cc)... 0 50 300 15 yes 0.03 0.45 0.01 0.35 0.14
select Pitch minfo
minfo = Get minimum... 0 0 Hertz parabolic
minfoFix\$ = fixed\$ (minfo, 2)
minfoFix = number (minfoFix\$)

select Sound maxfo
To Pitch (raw cc)... 0 75 1600 15 yes 0.03 0.45 0.01 0.35 0.14
select Pitch maxfo
maxfo = Get maximum... 0 0 Hertz parabolic
maxfoFix\$ = fixed\$ (maxfo, 2)
maxfoFix = number (maxfoFix\$)

```
forange = 12*log10(maxfo/minfo)/log10(2)
forangeFix$ = fixed$ (forange, 0)
forangeFix = number (forangeFix$)
```

#Intensity Range

```
select Sound softc
softdB = Get intensity (dB)
softdBFix$ = fixed$ (softdB, 2)
softdBFix = number (softdBFix$)
```

```
select Sound loudc
louddB = Get intensity (dB)
louddBFix$ = fixed$ (louddB, 2)
louddBFix = number (louddBFix$)
```

```
irange = louddBFix-softdBFix
irangeFix$ = fixed$ (irange, 2)
irangeFix = number (irangeFix$)
```

#Glide Spectrogram

```
select Sound glide
To Spectrogram: 0.03, 5000, 0.002, 20, "Gaussian"
```

#Song spectrogram

```
select Sound song
To Spectrogram: 0.03, 5000, 0.002, 20, "Gaussian"
```

#Soft Spectrogram

```
select Sound soft
To Spectrogram: 0.03, 5000, 0.002, 20, "Gaussian"
```

#Soft fo

```
select Sound soft
To Pitch (cc)... 0 75 15 yes 0.03 0.8 0.01 0.35 0.14 1300
meansofffo = Get mean... 0 0 Hertz
```

#Soft Intensity

```
select Sound softc
softdB = Get intensity (dB)
softdBFix$ = fixed$ (softdB, 2)
softdBFix = number (softdBFix$)
```

#Soft CPPS

```
select Sound softc
To PowerCepstrogram... 60.0 0.002 5000.0 50.0
softcpps = Get CPPS: "no", 0.01, 0.001, 60, 1000, 0.05, "Parabolic", 0.001, 0, "Straight", "Robust"
softcppsFix$ = fixed$ (softcpps, 2)
softcppsFix = number (softcppsFix$)
```

#Soft L1-L2

```
selectObject: "Sound softc"
To Ltas: 50
l1 = Get maximum: (meansoftfo-50), (meansoftfo+50), "none"
l2 = Get maximum: ((meansoftfo*2)-50), ((meansoftfo*2)+50), "none"
l1l2 = l1-l2
l1l2Fix$ = fixed$ (l1l2, 2)
l1l2Fix = number (l1l2Fix$)
```

#MPT

```
select Sound mpt
mpts = Get total duration
mptsFix$ = fixed$ (mpts, 2)
mptsFix = number (mptsFix$)
```

#Report Printing

```
Erase all
Select inner viewport: 0, 9.5, 0, 5.5
Axes: 0, 1, 0, 1
Solid line
Line width: 1
Black
Helvetica
Font size: 15
Select outer viewport: 0.5, 9, 0.5, 0.55
Viewport text: "Left", "Top", 0, "##Baseline Singing Voice Diagnostic Metrics#"
```

Erase all

Font size: 7

Select outer viewport: 0.5, 9, 0.5, 0.55

Viewport text: "Left", "Bottom", 0, "Baker, 2025"

Font size: 12

Select inner viewport: 0.5, 9, 0.5, 0.8

Text: 0.0, "Left", 0.5, "Half", "Baseline Acoustic Assessment for the Singing Voice"

Font size: 9

Select inner viewport: 0.5, 9, 0.8, 1

Axes: 0, 1, 0, 1

Paint rectangle: "{1,1,1}", 0, 1, 0, 1

Draw inner box

Colour: "{0,0,0}"

Line width: 2

Draw line: 0.25, 0, 0.25, 1

Draw line: 0.45, 0, 0.45, 1

Draw line: 0.65, 0, 0.65, 1

Draw line: 0.8, 0, 0.8, 1

Colour: "{0,0,0}"

Text: 0.01, "Left", 0.5, "Half", "Name: 'name\$'"

Text: 0.26, "Left", 0.5, "Half", "Identifier: 'identifier\$'"

Text: 0.46, "Left", 0.5, "Half", "Assessment: 'assessment_date\$'"

Text: 0.66, "Left", 0.5, "Half", "DoB: 'date_of_Birth\$'"

Text: 0.81, "Left", 0.5, "Half", "Sex: 'sex_if_relevant\$'"

endif

selectObject: "Spectrogram soft"

Select inner viewport: 0.5, 6, 3.5, 5.5

Axes: 0, 1, 0, 1

Paint rectangle: "{1,1,1}", 0, 1, 0, 1

Paint: 0, 0, 0, 5000, 100, "yes", 50, 6, 0, "no"

Draw inner box

One mark left: 0, "no", "yes", "no", "0"

One mark left: 5000, "no", "yes", "no", "5000"

Text left: "no", "Frequency (Hz)"

Text bottom: "no", "Time (s)"

Marks bottom: 2, "yes", "no", "no"

Select inner viewport: 0.5, 6, 3.35, 3.5

Text top: "no", "Narrowband Spectrogram Soft C4/C5 (.03s)"

Select inner viewport: 0.5, 6, 3.5, 5.5

Text top: "no", "##Note vibrato regularity, formant clusters, bifurcations, and subharmonics.##"

selectObject: "Ltas softc"
Select inner viewport: 6.5, 9, 3.5, 5.5
Axes: 0, 1, 0, 1
Paint rectangle: "{1,1,1}", 0, 1, 0, 1
Draw: 0, 6000, -20, 70, "no", "curve"
Draw inner box
One mark left: -20, "no", "yes", "no", "-20"
One mark left: 70, "no", "yes", "no", "70"
Text left: "no", "SPL (dB/Hz)"
Text bottom: "no", "Frequency (Hz)"
Marks bottom: 2, "yes", "no", "no"
Select inner viewport: 6.5, 9, 3.35, 3.5
Text top: "no", "LTAS Soft C4/C5 (50Hz)"
Select inner viewport: 6.5, 9, 3.5, 5.5
Text top: "no", "##Note harmonic strength and high-frequency noise.##"

selectObject: "Spectrogram glide"
Select inner viewport: 0.5, 9, 6.25, 8.25
Axes: 0, 1, 0, 1
Paint rectangle: "{1,1,1}", 0, 1, 0, 1
Paint: 0, 0, 0, 5000, 100, "yes", 50, 6, 0, "no"
Draw inner box
One mark left: 0, "no", "yes", "no", "0"
One mark left: 5000, "no", "yes", "no", "5000"
Text left: "no", "Frequency (Hz)"
Text bottom: "no", "Time (s)"
Marks bottom: 2, "yes", "no", "no"
Select inner viewport: 0.5, 9, 6.1, 8
Text top: "no", "Narrowband Spectrogram Glide (.03s)"
Select inner viewport: 0.5, 9, 6.25, 8.25
Text top: "no", "##Note registral transitions and phonation gaps.##"

selectObject: "Spectrogram song"
Select inner viewport: 0.5, 9, 9, 11
Axes: 0, 1, 0, 1
Paint rectangle: "{1,1,1}", 0, 1, 0, 1
Paint: 0, 0, 0, 5000, 100, "yes", 50, 6, 0, "no"
Draw inner box
One mark left: 0, "no", "yes", "no", "0"
One mark left: 5000, "no", "yes", "no", "5000"
Text left: "no", "Frequency (Hz)"
Text bottom: "no", "Time (s)"
Marks bottom: 2, "yes", "no", "no"

Select inner viewport: 0.5, 9, 8.85, 10.75

Text top: “no”, “Narrowband Spectrogram Song (.03s)”

Select inner viewport: 0.5, 9, 9, 11

Text top: “no”, “##Note harmonicity, presence of noise or subharmonics, and onset/offset quality.##”

Font size: 10

Select inner viewport: 0.5, 4, 1, 1.25

Axes: 0, 1, 0, 1

Text: 0.01, “Left”, 0.5, “Half”, “SVHI-10: ##’svhi10’#”

Select inner viewport: 0.5, 4, 1.25, 1.5

Text: 0.01, “Left”, 0.5, “Half”, “EASE VF: ##’easevf’/40#”

Line width: 2

Draw line: 0, 1, 1, 1

Line width: 1

Draw line: 0, 0, 1, 0

Select inner viewport: 0.5, 4, 1.5, 1.75

Text: 0.01, “Left”, 0.5, “Half”, “EASE PRI: ##’easepri’/40#”

Line width: 1

Draw line: 0, 0, 1, 0

Select inner viewport: 0.5, 4, 1.75, 2

Text: 0.01, “Left”, 0.5, “Half”, “EASE VC: ##’easevc’/8#”

Line width: 2

Draw line: 0, 0, 1, 0

Select inner viewport: 0.5, 4, 2, 2.25

Text: 0.01, “Left”, 0.5, “Half”, “Minimum fo: ##’minfoFix’ Hz#”

Line width: 2

Draw line: 0, 0, 1, 0

Select inner viewport: 0.5, 4, 2.25, 2.5

Text: 0.01, “Left”, 0.5, “Half”, “Max fo: ##’maxfoFix’ Hz#”

Line width: 1

Draw line: 0, 0, 1, 0

Select inner viewport: 0.5, 4, 2.5, 2.75

Text: 0.01, “Left”, 0.5, “Half”, “Frequency Range: ##’forangeFix’ Semitones##”

Line width: 2

Draw line: 0, 0, 1, 0

Select inner viewport: 5.5, 9, 1, 1.25

Text: 0.01, “Left”, 0.5, “Half”, “Calibration Factor: ##’calfacFix’ dB#”

Select inner viewport: 5.5, 9, 1.25, 1.5

Text: 0.01, “Left”, 0.5, “Half”, “SPL C4/C5 Soft: ##’softdBFix’ dB#”

Line width: 2

Draw line: 0, 1, 1, 1

Line width: 1

Draw line: 0, 0, 1, 0
Select inner viewport: 5.5, 9, 1.5, 1.75
Text: 0.01, "Left", 0.5, "Half", "SPL C4/C5 Loud: ##'louddBFix' dB#"
Line width: 1
Draw line: 0, 0, 1, 0
Select inner viewport: 5.5, 9, 1.75, 2
Text: 0.01, "Left", 0.5, "Half", "SPL C4/C5 Range: ##'irangeFix' dB#"
Line width: 2
Draw line: 0, 0, 1, 0
Line width: 2
Select inner viewport: 5.5, 9, 2, 2.25
Text: 0.01, "Left", 0.5, "Half", "Soft C4/C5 CPPS: ##'softcppsFix' dB#"
Draw line: 0, 0, 1, 0
Select inner viewport: 5.5, 9, 2.25, 2.5
Text: 0.01, "Left", 0.5, "Half", "L1-L2 C4/C5 : ##'l1l2Fix' dB#"
Draw line: 0, 0, 1, 0
Select inner viewport: 5.5, 9, 2.5, 2.75
Text: 0.01, "Left", 0.5, "Half", "MPT C4/C5: ##'mptsFix' Seconds##"
Draw line: 0, 0, 1, 0

Copy Praat picture

Select inner viewport: 0.5, 9, 11.5, 11.75
Text: 0, "Left", 0, "Half", "##Comments:# 'Include_perceptual_features_or_data_from_other_measures_below\$"
Select inner viewport: 0.5, 9, 0.5, 12
Copy to clipboard

select all
minus Sound minfo
minus Sound maxfo
minus Sound glide
minus Sound soft
minus Sound loud
minus Sound mpt
minus Sound song
Remove