
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Mittapalle, Kiran; Pohjalainen, Hilla; Helkkula, Pyry; Kaitue, Kasimir; Minkkinen, Mikko; Tolppanen, Heli; Nieminen, Tuomo; Alku, Paavo

Glottal flow characteristics in vowels produced by speakers with heart failure

Published in:
Speech Communication

DOI:
[10.1016/j.specom.2021.12.001](https://doi.org/10.1016/j.specom.2021.12.001)

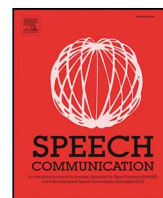
Published: 01/02/2022

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Mittapalle, K., Pohjalainen, H., Helkkula, P., Kaitue, K., Minkkinen, M., Tolppanen, H., Nieminen, T., & Alku, P. (2022). Glottal ow characteristics in vowels produced by speakers with heart failure. *Speech Communication*, 137, 35-43. <https://doi.org/10.1016/j.specom.2021.12.001>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.



Glottal flow characteristics in vowels produced by speakers with heart failure

Mittapalle Kiran Reddy ^{a,*}, Hilla Pohjalainen ^a, Pyry Helkkula ^b, Kasimir Kaitue ^a,
Mikko Minkkinen ^c, Heli Tolppanen ^c, Tuomo Nieminen ^d, Paavo Alku ^a

^a Department of Signal Processing and Acoustics, Aalto University, 00076 Aalto, Finland

^b Institute of Molecular Medicine, University of Helsinki, 00290 Helsinki, Finland

^c Heart and Lung Center, Helsinki University Hospital, 00290 Helsinki, Finland

^d Päijät-Häme Joint Authority for Health and Wellbeing, Lahti, Finland

ARTICLE INFO

Keywords:

Heart failure
Glottal source parameters
Sound pressure level
Wilcoxon rank sum test
Bio-marker

ABSTRACT

Heart failure (HF) is one of the most life-threatening diseases globally. HF is an under-diagnosed condition, and more screening tools are needed to detect it. A few recent studies have suggested that HF also affects the functioning of the speech production mechanism by causing generation of edema in the vocal folds and by impairing the lung function. It has not yet been studied whether these possible effects of HF on the speech production mechanism are large enough to cause acoustically measurable differences to distinguish speech produced in HF from that produced by healthy speakers. Therefore, the goal of the present study was to compare speech production between HF patients and healthy controls by focusing on the excitation signal generated at the level of the vocal folds, the glottal flow. The glottal flow was computed from speech using the quasi-closed phase glottal inverse filtering method and the estimated flow was parameterized with 12 glottal parameters. The sound pressure level (SPL) was measured from speech as an additional parameter. The statistical analyses conducted on the parameters indicated that most of the glottal parameters and SPL were significantly different between the HF patients and healthy controls. The results showed that the HF patients generally produced a more rounded glottal pulse and a lower SPL level compared to the healthy controls, indicating incomplete glottal closure and inappropriate leakage of air through the glottis. The results observed in this preliminary study indicate that glottal features are capable of distinguishing speakers with HF from healthy controls. Therefore, the study suggests that glottal features constitute a potential feature extraction approach which should be taken into account in future large-scale investigations in studying the automatic detection of HF from speech.

1. Introduction

Heart failure (HF) is one of the leading causes of mortality worldwide, affecting the lives of both patients and their caregivers (Savarese and Lund, 2017). It is a chronic, progressive condition in which the heart muscle is unable to pump enough blood to meet the body tissue's needs for metabolism (Coronel et al., 2001; Ponikowski et al., 2016). Traditionally, HF is broadly classified according to the left ventricular ejection fraction (LVEF) into two categories: (1) HF with reduced ejection fraction (Murphy et al., 2020) (LVEF \leq 40), where the heart muscle is “weak”, and (2) HF with preserved ejection fraction (Vasan et al., 2018) (LVEF $>$ 40), where the heart muscle becomes “stiff”. The conditions that may lead to HF include: cardiomyopathy, coronary artery disease (CAD), emphysema, diabetes and so on. In spite of the etiology and mechanism (including ejection fraction), common signs and symptoms of HF include, for example, edema, dyspnea, dizziness, irregular heart beats, and hypo-perfusion. Recently, the question

concerning the impact of HF on the acoustic characteristics of speech signals has awakened interest among speech scientists in investigating the potential use of voice as bio-marker for heart diseases (Maor et al., 2016, 2020; Murton et al., 2017; Sara et al., 2020; Kiran Reddy et al., 2021). Voice has the potential to provide an easily obtained, non-invasive way to monitor physiological changes throughout the body, as long as those changes also affect the larynx (Murton et al., 2017). The larynx is a component of the respiratory tract, and has several important functions, including phonation. Some studies have reported a consistent association between phonation, which provides the periodic excitation for voiced speech, and the cardiovascular system (Orlikoff and Baken, 1989; Alvear et al., 2013; Murton et al., 2017; Maor et al., 2016; Sara et al., 2020). The heartbeat was shown to modulate the fundamental frequency (F_0) of voice signals in Orlikoff and Baken (1989).

* Corresponding author.

E-mail address: kiran.r.mittapalle@aalto.fi (K.R. Mittapalle).

<https://doi.org/10.1016/j.specom.2021.12.001>

Received 11 September 2021; Received in revised form 18 November 2021; Accepted 2 December 2021

Available online 5 January 2022

0167-6393/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Their work concluded that changes in F_0 were caused by pressure-related changes in the stiffness of the vocal fold vascular bed and by heartbeat cyclic alterations of the geometry of the thyro-arytenoid muscle. The study by authors in [Alvear et al. \(2013\)](#) corroborates the findings in [Orlikoff and Baken \(1989\)](#), and adds complementary analysis to make evident that all cardiovascular parameters (such as heartbeat and blood pressure) are associated with the laryngeal vocal function. More recently, in [Murton et al. \(2017\)](#), it has been hypothesized that the HF-related edema in the vocal folds and lungs affect phonation and speech respiration. The authors analyzed the voices of HF patients as they underwent treatment for decompensated HF and returned to a stable clinical state. Their results demonstrated that patients spoke faster or breathed less frequently after HF treatment and most patients showed increased irregularity and decreased F_0 at admission compared to discharge. Furthermore, the authors emphasize that a small increase in the amount of edema leads to measurable changes in the voice characteristics. Using acoustic features such as mel cepstrum, pitch, jitter, shimmer and loudness, the authors in [Maor et al. \(2016, 2020\)](#) showed that vocal bio-markers are associated with increased risk of hospitalization and mortality among HF patients. These authors speculated that the vagus nerve, which is one of the most important cranial nerves for speech production and is also critical for autonomic control of the heart, can be a possible link between HF and voice. Although the previous studies referred to above have reported an association between HF and voice production, a detailed understanding of the behavior of the voice source in speakers with HF is lacking.

Phonation is brought about by air flowing from the lungs, which results in periodic vibration of the vocal folds. The vocal folds are a pair of layer-structured tissues located in the larynx, and the gap between the two vocal folds is known as the glottis. The source of (voiced) speech is the modulated airflow generated by the vibrating vocal folds, the glottal volume velocity waveform (also known as the glottal flow) ([Fant, 1970](#)). Speech is produced as a result of filtering the glottal flow by the vocal tract (i.e., the vocal cavity whose shape and dimensions are modulated by movements of, for example, the tongue and the lips) and by converting the glottal flow signal at the mouth into a pressure signal in free field (i.e., the lip radiation effect). Many previous studies have shown that the shape of the glottal flow waveform is strongly correlated with the subglottal pressure and vocal fold vibration pattern ([Sundberg et al., 1993](#); [Vilkman et al., 2002](#); [Holmberg et al., 1988](#); [Gauffin and Sundberg, 1989](#); [Alku et al., 2006a](#)). In HF patients, variations in the subglottal pressure and vocal fold vibration patterns may occur due to the presence of edema or other factors. Edema is the hallmark of HF symptoms. Depending on the HF mechanism, generation of edema can be more prominent in the systemic circulation (outside the lungs) and/or in the pulmonary circulation (in the lungs). At a stable stage of HF, edema can be absent, but it progressively increases with the deterioration of the disease. The presence of edema in the lung tissue can cause illnesses such as dyspnea (shortness of breath), dizziness, and chest pain in HF patients ([Hecht, 1956](#)). This can be expected to reduce the subglottal pressure, thereby also affecting the shape of the glottal flow pulse. If edema mainly occurs in the systemic circulation, the patient may not have edema in the lungs at all. However, the vocal folds can be affected in that case too because the vocal folds consist of thin tissue layers that might be particularly sensitive to the HF-related edema. For example, in [Verdolini et al. \(2002\)](#) it has been found that a dose of diuretic Lasix (furosemide), which is widely used to treat decompensated HF, induced a 23% increase in the phonation threshold pressure in healthy adults. Furthermore, left vocal cord paralysis is frequently observed with various forms of intrathoracic diseases ([William and Diefenbach, 1949](#)). Because both the impairment of lung function and vocal fold movements result in anomalous glottal flows, it can be hypothesized that HF patients may show different characteristics of the glottal flow pulse compared to healthy controls.

The aim of this study was to first investigate the behavior of the source of speech generated at the level of the vocal folds, the glottal

Table 1
General information of the HF patients.

Age (mean, standard deviation)		67, 13
Male sex		70%
Etiology of HF	Ischemic cardiomyopathy	35%
	Dilated cardiomyopathy	25%
	Tachycardia myopathy	15%
	Valvular heart disease	15%
	Other	10%
Ejection fraction	>40%	35%
	≤40%	65%
Diagnosed clinical edema		50%
Diagnosed lung disease		10%
Medication	Loop diuretic medication	100%
	B blocker medication	85%

flow, in HF patients, and then to compare the differences between glottal flows produced by HF patients and their healthy controls. In several previous studies, the glottal flow analysis has been shown to be very useful in understanding the voice source characteristics in abnormal and disordered voices ([Gillespie et al., 2017](#); [Drugman et al., 2014](#); [Sundberg et al., 1993](#); [Sapienza et al., 1998](#); [Alku, 2011](#); [Holmberg et al., 1988](#)). The glottal flow was estimated from microphone speech signals produced by 20 HF patients and by 25 healthy talkers using a non-invasive inversion technique called glottal inverse filtering (GIF) ([Alku, 2011](#); [Airaksinen et al., 2013](#)). The estimated glottal flow pulses were expressed using time- and frequency-domain parameters and statistical analyses were conducted for the computed parameters to investigate whether the parameter distributions are different between the two speaker groups. SPL was used in addition to the glottal parameters to compare speech signals produced by the two speaker groups. SPL is a widely used parameter of speech production whose value is regulated primarily by means of subglottal pressure ([Sundberg et al., 1993](#)). In addition, SPL has been shown to be related to the shape of the glottal flow waveform ([Vilkman et al., 2002](#); [Holmberg et al., 1988](#)). To the best of our knowledge, this is the first study in which the glottal flow is analyzed in HF. In addition to providing a specific insight into the glottal flow characteristics in HF, the study also extends the currently scant knowledge about the characteristics of speech signals in HF patients.

2. Material and methods

2.1. Dataset

Currently, there are no open source databases available to study how heart diseases affect acoustic characteristics of speech signals. Therefore, a new speech dataset was recorded in Finnish by the authors as a part of the current study. Currently, this database includes speech recordings of 20 HF patients (14 males, 6 females) and 25 healthy controls (18 males, 7 females). The ages of the HF patients vary between 36 years and 81 years (mean: 67 years, standard deviation: 13 years) and the ages of the healthy speakers vary between 54 years and 83 years (mean: 60 years, standard deviation: 7 years). In the patient group, only one patient's age is below 50 years (36 years). The average age of onset for HF differs between males and females, and certain symptoms like shortness of breath are more common in female patients. However, there are no differences in treatment for the male and female HF patients of the current study. The general information about the HF speakers is provided in [Table 1](#). In the table, the age information of the patients is shown in the first row and the rest of the rows show other information as the percentage of patients in each category or sub-category listed in the first and second columns. The information about the two issues (existence of edema and diagnosed lung disease) was deliberately collected from the patients because these issues have been associated in previous studies ([Hecht, 1956](#); [Murton et al., 2017](#)) with production of speech in HF. [Table 1](#) shows that only 50% of the patients

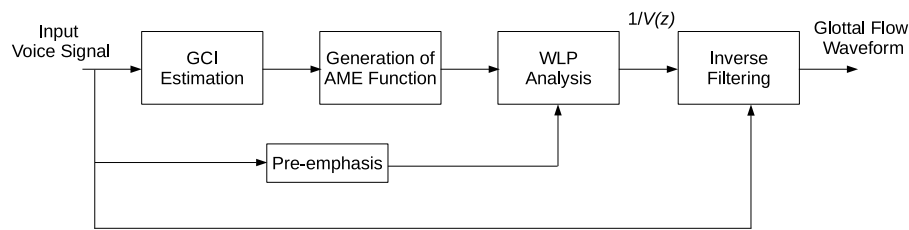


Fig. 1. The block diagram of the QCP glottal inverse filtering method (redrawn from Airaksinen et al. (2013)). Glottal closure instants (GCIs) are first estimated from the input voice signal and then used to generate the attenuated main excitation (AME) function. The AME function is used to derive the weighted linear prediction (WLP)-based vocal tract model $1/V(z)$ using the pre-emphasized voice signal. Finally, the estimate of the glottal flow waveform is obtained by inverse filtering the input signal with $V(z)$.

had clinically observable edema but the table also indicates that all the 20 patients took loop diuretic, which is a medicine for edema. Edema is in fact a broad concept, and even if there is no clinically detectable edema in the pulmonary and systemic circulation, there may be edema in the blood vessels and at the tissue level (which might affect voice). The patients in this study were hospitalized for HF of any etiology, regardless of the left ventricular ejection fraction, either due to acute worsening of the symptoms or for diagnostic tests due to advanced HF. The majority of the patients in the current study had HF with reduced ejection fraction and remaining patients had preserved ejection fraction.

Each speaker read the same Finnish text three times (the text reading task) and produced one piece of spontaneous speech. All the speakers were L1 speakers of Finnish. The text in the text reading task was a brief imaginary weather report consisting of three paragraphs and including altogether 91 words. Because the goal of the study was to compare the glottal flow characteristics between the HF patients and healthy controls, the acoustic analyses focused on three selected segments of voiced speech in the recorded speech signal of the middle recitation of the text reading task. All of these three segments included the vowel [a:], which is the vowel with the highest first formant (F1) among the Finnish vowels. The vowel [a] was selected because it is known that the estimation of the glottal flow with GIF (to be described in Section 2.2) is most accurate for this vowel due to the high value of its F1 (Alku et al., 2006b). GIF analysis can, however, be computed in principle for any voiced segment of speech by using automatic inverse filtering and parameterization as done, for example, in studies where GIF has been used in speech synthesis (Airaksinen et al., 2018; Cui et al., 2018). For some phones, the accuracy of GIF might be lower compared to [a] due to nasal coupling (for the nasal [ŋ]) or due to poor separation of the source and tract when the phone has low F1 (e.g. the vowel [i]). Therefore, in order to maximize the accuracy of GIF in this preliminary study, we considered it justified to focus only on the vowel [a].

In order to neutralize the potential effect of the phonetic stress in vowel productions, the three segments were extracted from the recorded speech in positions that corresponded to the first syllable of the first word in each of the three text paragraphs. In all three segments, the vowel was preceded and followed by oral consonants so there was no anticipatory or carryover nasalization. Hence, the total amount of speech data to be analyzed consisted of 60 utterances produced by the HF patients ($14 \times 3 = 42$ by males, and $6 \times 3 = 18$ by females) and of 75 utterances produced by the healthy controls ($18 \times 3 = 54$ by males, and $7 \times 3 = 21$ by females). The speech data, sampled at 44.1 kHz, were recorded in doctor's practice room using a headset condenser microphone (DPA 4065-BL) and an AD converter (RME Babyface Pro). The mouth-to-microphone distance was 5 cm. In order to estimate the SPL of the produced vowels, a calibration tone of 1 kHz with a constant SPL of 94 dB was recorded using a sound calibrator (Amprobe SM-CAL1). On the computer, the speech signals were first down-sampled to 16 kHz and then high-pass filtered with a linear phase FIR filter (cut-off frequency: 60 Hz) to remove the low-frequency noise picked up by the recording microphone.

2.2. Glottal inverse filtering

Glottal inverse filtering (GIF) refers to the technique of estimating the glottal flow excitation from the speech (pressure) signal recorded by microphone. Several GIF methods have been proposed in the literature (for a review on GIF, see Alku (2011)). In this study, we utilized the quasi-closed phase (QCP) algorithm proposed in Airaksinen et al. (2013) as the GIF method. The use of QCP as the GIF algorithm in the current study is justified on the grounds that it was shown in Airaksinen et al. (2013) to perform better than four state-of-the-art GIF methods for both modal and non-modal phonation types. The block diagram in Fig. 1 describes the steps involved in the QCP method. The method is based on the principles of closed phase (CP) analysis (Wong et al., 1979), which estimates the vocal tract model from a few speech samples located in the CP of the glottal cycle using linear prediction (LP) analysis. Unlike CP analysis, QCP takes advantage of all the speech samples of the analysis frame in computing the vocal tract model. This is enabled by using weighted linear prediction (WLP) analysis with an attenuated main excitation (AME) (Alku et al., 2013; Airaksinen et al., 2013) weighting function as an all-pole modeling method in the estimation of the vocal tract. As shown in the studies by Alku et al. (2013) and Airaksinen et al. (2013), the use of WLP together with the AME function aims at downgrading the prominent effect of the glottal source in the computation of the vocal tract all-pole model. This downgrading corresponds to attenuating the prediction error energy in all-pole modeling at the instants of glottal closure where the effect of the glottal excitation is strong. Consequently, the resulting vocal tract model (denoted by $1/V(z)$ in Fig. 1) is affected more by the characteristics of the vocal tract resonances, leading to less biasing of the tract model by the glottal source. According to Alku et al. (2013), Airaksinen et al. (2013), the AME weighting function is a simple, real-valued time-domain signal whose values are equal to 1.0 in the samples during the glottal open phase and equal to a small positive value (e.g., 0.03) in the vicinity of glottal closure. The AME waveform has three parameters (the duration quotient, the position quotient and the value of the waveform at glottal closure), whose values were set in the current study as in Airaksinen et al. (2013). In addition, the generation of the AME waveform calls for extracting glottal closure instants (GCIs). The GCIs were estimated using the recently proposed continuous wavelet transform-based approach developed in Keerthana et al. (2019). Once the vocal tract model has been computed with WLP, the input acoustic speech signal is finally filtered in the QCP method with the inverse of the computed vocal tract transfer function to obtain the estimate of the glottal flow waveform.

2.3. Parameterization

The glottal flow waveforms estimated with QCP were parameterized using a glottal parameter set consisting of 12 known time- and frequency-domain parameters (Alku et al., 2002; Childers and Lee, 1991). These parameters characterize various aspects of the glottal flow waveform, and were estimated using the APARAT toolbox (Airas et al., 2005). The glottal parameters are listed in Table 2. The open

Table 2

Time- and frequency-domain glottal parameters. For more details, see (Alku et al., 2002; Childers and Lee, 1991).

Time-domain glottal parameters	
OQ1	Open quotient, calculated from the primary glottal opening
OQ2	Open quotient, calculated from the secondary glottal opening
OQa	Open quotient, derived from the LF model
QOQ	Quasi-open quotient
ClQ	Closing quotient
SQ1	Speed quotient, calculated from the primary glottal opening
SQ2	Speed quotient, calculated from the secondary glottal opening
AQ	Amplitude quotient
NAQ	Normalized Amplitude quotient
Frequency-domain glottal parameters	
HRF	Harmonic richness factor
PSP	Parabolic spectral parameter
H1H2	Difference between first two glottal harmonics

quotient (OQ) is defined as the ratio of the duration of the glottal open phase and the duration of the glottal cycle (Sapienza et al., 1998). The quasi-open quotient (QoQ) measures the duration of the open phase as a time span between the instant when the ascending glottal flow crosses a level that is 50% of the maximum AC flow and the instant when the descending flow falls below this level (Gillespie et al., 2017). The closing quotient (ClQ) is defined as the ratio between the duration of the closing phase and the duration of the glottal cycle (Monsen and Engebretson, 1977). The speed quotient (SQ) is defined as the ratio between the duration of the opening phase and the duration of the closing phase (Sapienza et al., 1998). The amplitude quotient (AQ) (Alku and Vilkmán, 1996) and the normalized amplitude quotient (NAQ) (Alku et al., 2002) are robust glottal source parameters which are based on quantifying time-domain properties of the glottal closing phase using two amplitude-domain measures extracted from the glottal flow and its derivative. As shown in Alku and Vilkmán (1996) and Alku et al. (2002), these two amplitude-domain measures are the AC amplitude of the glottal flow and the amplitude of the minimum of the flow derivative. AQ parameterizes the length of the glottal closing phase as the ratio of these two amplitude values and NAQ normalizes this ratio by the length of the glottal cycle. HRF measures the spectral tilt of the glottal source as the ratio between the energy of the fundamental and the sum of the energies of the higher harmonics in the glottal flow spectrum (Childers and Lee, 1991). PSP measures the spectral tilt by fitting a parabola to low frequencies of the glottal source spectrum (Alku et al., 1997). Finally, H1H2 is a straightforward measure for the spectral tilt of the glottal source defined as the dB-difference between the levels of the first two harmonics, H1 and H2, of the glottal source spectrum (Fant, 1970). The parameters described above have been used in many studies investigating glottal source characteristics in speech and voice signals produced by healthy speakers (e.g., (Alku et al., 2006b; Vilkmán et al., 2002)). In addition, many of these parameters have been used in the analysis and detection of disorders such as Parkinson's disease (Novotný et al., 2020; Narendra et al., 2021), vocal hyperfunction (Espinoza et al., 2017) and dysarthria (Gillespie et al., 2017) from speech signals.

The glottal parameters were computed in 30 ms frames using a frame shift of 10 ms. While HRF and H1H2 were computed pitch-asynchronously once per frame, the remaining parameters were computed pitch-synchronously once per glottal cycle and then averaged over the frame. All nine time-domain parameters and the PSP were expressed using a linear scale while H1H2 and HRF were expressed using the dB scale. The parameters computed from all frames of the input speech signal were averaged over the utterance. In addition to the glottal parameters, the SPL was also computed from the vowel segments analyzed. The SPL was estimated using the approach described in Švec and Granqvist (2018) by comparing the average of the square of the speech waveform to that of the calibration tone (described in Section 2.1). The SPL measure is closely related to shortness of breath, a HF symptom more commonly observed in female patients.

Table 3

The p -values of the Wilcoxon rank sum test for each of the 12 glottal parameters and for SPL in comparing the two speaker groups (HF patients vs. healthy speakers). The tests were conducted separately for male (M) and female (F) speakers. The statistically significant p -values are shown in bold. The values within parentheses denote the z-score.

Parameter	p -value (M)	p -value (F)
OQ1	0.000000 (5.43873)	0.199907 (1.281817)
OQ2	0.165827 (1.38573)	0.303822 (1.028271)
OQa	0.007825 (2.65953)	0.099342 (1.648051)
QOQ	0.000000 (5.24573)	0.190200 (1.309989)
ClQ	0.000095 (3.90245)	0.047021 (−1.986112)
SQ1	0.000249 (−3.66313)	0.014815 (2.436861)
SQ2	0.014550 (−2.44337)	0.027002 (2.211487)
AQ	0.000011 (4.38881)	0.490063 (0.690209)
NAQ	0.010261 (2.56689)	0.405935 (−0.831068)
HRF	0.000013 (4.35793)	0.359885 (0.915584)
PSP	0.019934 (2.32757)	0.374857 (0.887412)
H1H2	0.000014 (−4.34249)	0.526168 (−0.633866)
SPL	0.015573 (−2.418770)	0.033704 (−1.982566)

3. Results

This section describes the results of the statistical tests that were computed from the glottal parameters and SPL in order to study whether these parameters show significant differences between the two speaker groups (the speakers with HF vs. the healthy controls). It should be noted that the two groups are similar in terms of age and none of the speakers had any known disorders (e.g., a common cold) that could have affected their speech production at the time of the recordings, except for HF in the case of the HF group. In addition, the language and the phonetic context of the speech sounds studied were the same for all the speakers. Therefore, it can be assumed that if the statistical tests showed a significant difference between the two speaker groups, the difference can be expected to be due to HF or confounding factors of HF status.

Statistical tests were carried out with the Wilcoxon rank sum test (Hollander et al., 2013) to compare the glottal flow characteristics and SPL in speech signals produced by the HF patients and healthy controls. The Wilcoxon rank sum test was chosen as it does not require any assumptions about the shape of the distribution and is effective even for small sample sizes (Morgan, 2017). The parameters computed from all the frames of an utterance were first averaged to obtain the utterance-level parameters for which the statistical tests were computed. The results of the Wilcoxon tests are given in Table 3. From the table, it can be observed that all the parameters (except for OQ2) showed statistically significant ($p < 0.05$) differences between the HF patients and the healthy controls for the male speakers. On the other hand, for the female speakers, only four parameters (ClQ, SQ1, SQ2 and SPL) were found to differ significantly.

In addition to the Wilcoxon test, the statistical distributions of the parameters between the two speaker groups are described using box plots in Figs. 2, 3, 4 and 5. Figs. 2 and 3 show the distributions of all the parameters for the male speakers. The observations from the box plots of the male speakers are as follows: The box plots demonstrate that the glottal parameters discriminate the HF patients from the healthy controls. In addition, these box plots describe the general differences in the glottal flow pulse between the HF patients and the healthy controls as follows. Three time-domain parameters shown in Fig. 2 demonstrate that the mean values of the two open quotients (OQ1 and OQ2) and the mean value of the closing quotient (ClQ) are larger for the HF patients compared to the healthy controls. Moreover, the box plots show that the mean values of the two speed quotients (SQ1 and SQ2) are closer to 1.0 for the HF patients compared to the healthy controls. As described in previous studies on the parameterization of human speech production (Alku, 2011; Holmberg et al., 1988), an increase in open quotients and ClQ combined with the value of speed quotients moving close to 1.0 implies that the glottal flow pulse changes to a more

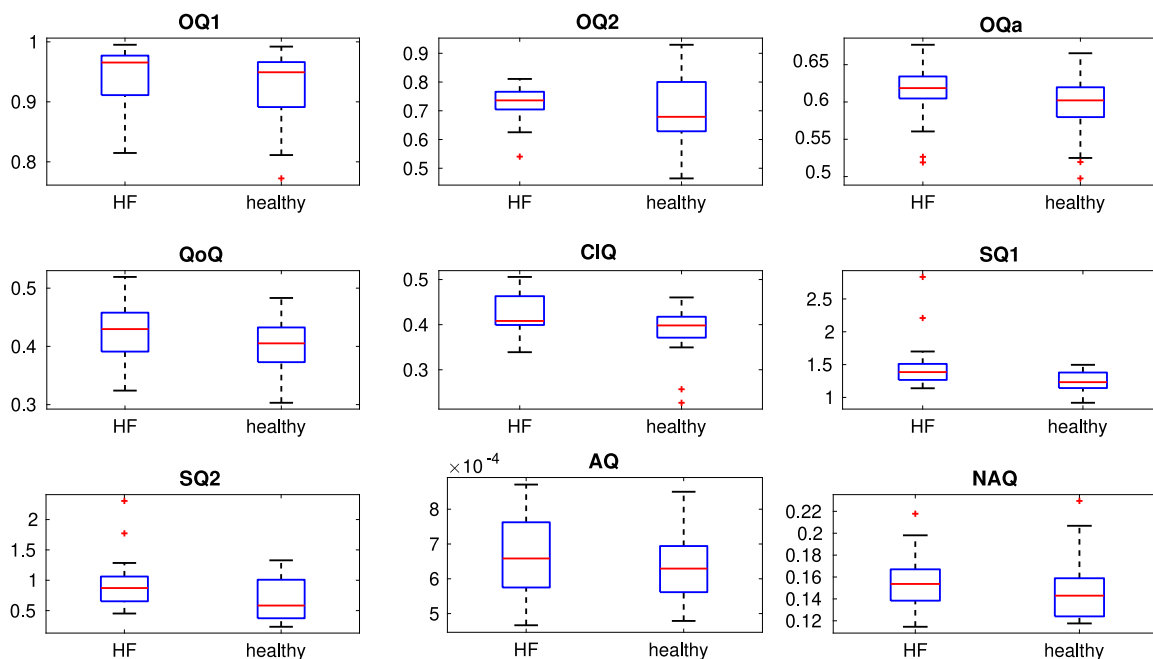


Fig. 2. The box plots of the time-domain glottal parameters computed for the male HF patients and for the male healthy controls. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles respectively. The whiskers on either side cover all points that are within 1.5 times the interquartile range, and points beyond these whiskers are plotted as outliers using the “+” symbol.

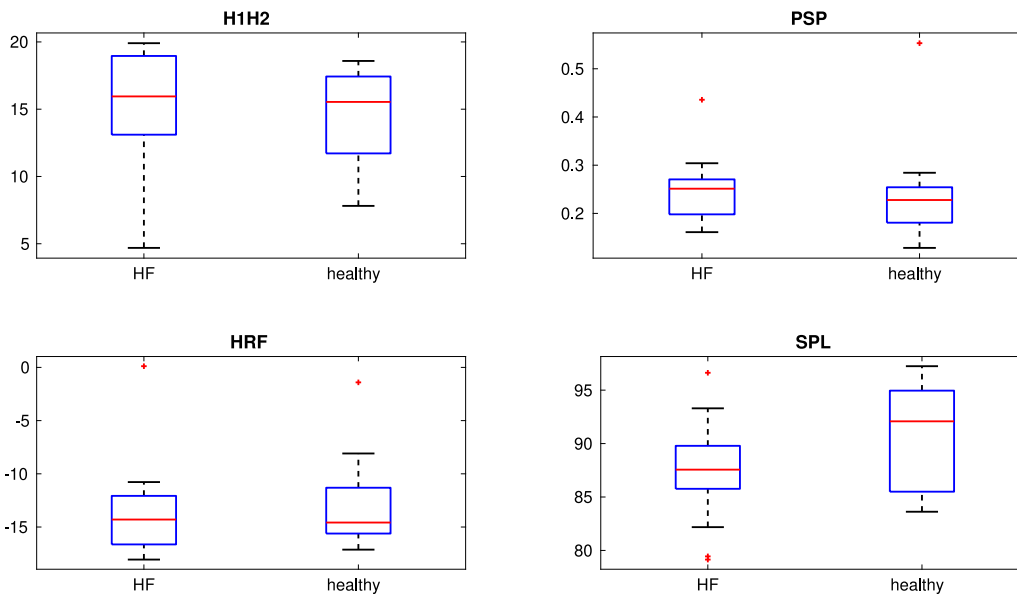


Fig. 3. The box plots of the frequency-domain glottal parameters and the box plot for SPL computed for the male HF patients and for the male healthy controls. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles respectively. The whiskers on either side cover all points that are within 1.5 times the interquartile range, and points beyond these whiskers are plotted as outliers using the “+” symbol.

rounded, soft shape. This implies that the pulse shape is more symmetric in the open phase and lacks sharp, abrupt changes in its waveform during the closing phase. The mean value of H1H2 is higher for the HF speakers compared to the healthy speakers, indicating incomplete glottal closure associated with breathy voice characteristics (Thompson et al., 2011). On the other hand, HRF values are lower for the HF patients, indicating a weaker harmonic structure in their glottal source spectrum. Both AQ and NAQ (AQ normalized with the length of the glottal cycle) reflect the behavior of the glottal flow during the closing phase of the glottal cycle (Vilkman et al., 2002). The mean AQ and NAQ values are higher for the male speakers with HF, indicating a more symmetrical glottal flow waveform. Higher mean values of AQ,

NAQ and PSP indicate that the male HF speakers tended to produce a breathier type of phonation (Alku et al., 1997; Vilkman et al., 2002). Lower values of the mean SPL indicate that the male HF patients generated on average softer speech compared to the healthy controls. The lower SPL may be attributed to incomplete glottal closure and the use of decreased subglottal pressure values (Sundberg et al., 1993; Alku, 2011; Holmberg et al., 1988). Subglottal pressure corresponds closely to lung pressure (Titze, 1992) and it can be estimated from oral-pressure measurements during production of stop consonants (typically /p/) using a pressure transducer connected to the speaker’s mouth with a plastic catheter (for more details, see, e.g., Alku et al. (2006a)). In a previous study (Alku et al., 2006a), it was shown that there is a very

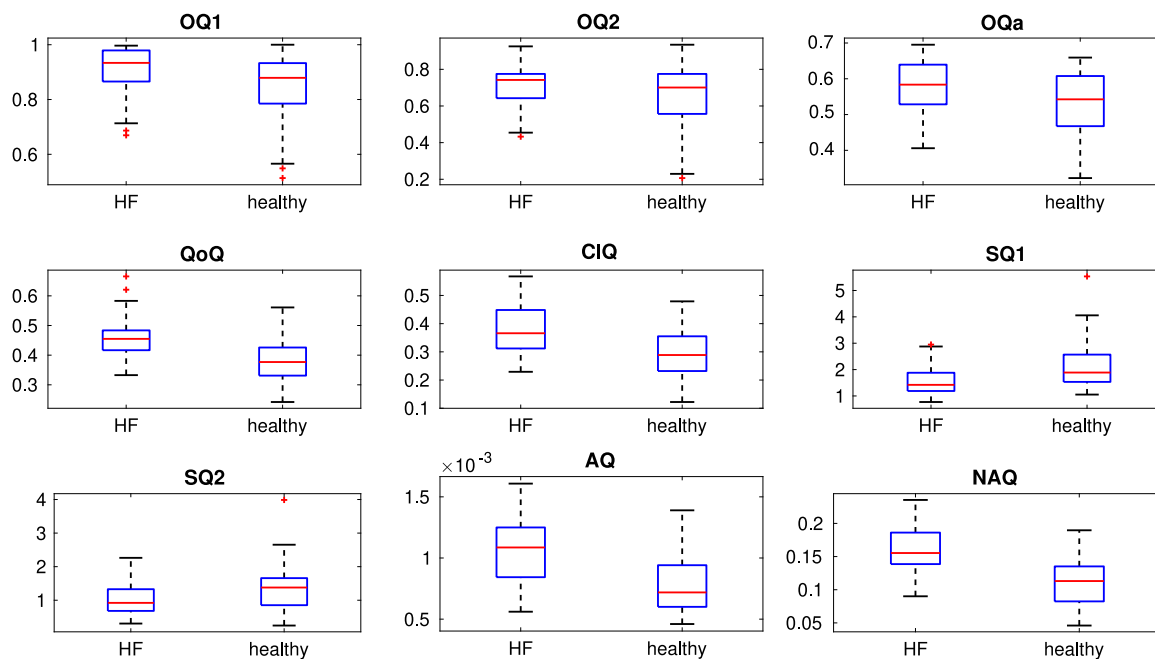


Fig. 4. The box plots of the time-domain glottal parameters computed for the female HF patients and for the female healthy controls. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles respectively. The whiskers on either side cover all points that are within 1.5 times the inter-quartile range, and points beyond these whiskers are plotted as outliers using the “+” symbol.

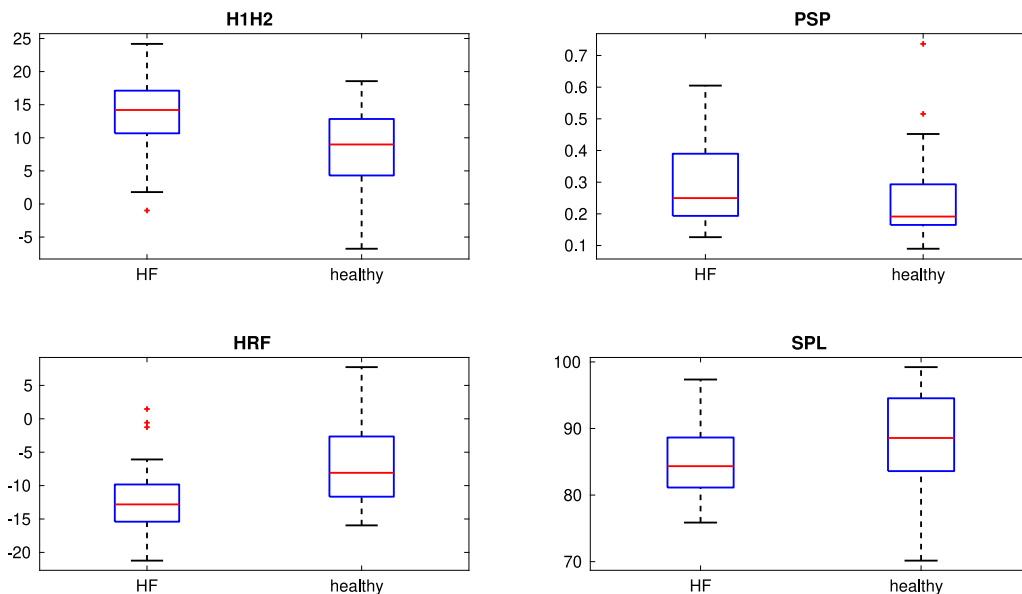


Fig. 5. The box plots of the frequency-domain glottal parameters and the box plot for SPL computed for the female HF patients and for the female healthy controls. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles respectively. The whiskers on either side cover all points that are within 1.5 times the inter-quartile range, and points beyond these whiskers are plotted as outliers using the “+” symbol.

large correlation between the mean log-transformed subglottal pressure and SPL. Given this previously reported (almost) linear correspondence between subglottal pressure and SPL and given the fact that the HF patients of the current study showed significantly lower levels of SPL, it can be argued that the HF patients have produced speech with lower lung pressure compared to the controls. Interestingly, however, none of the HF patients (male and female) were diagnosed with chronic lung disease (see Table 1). It is therefore possible that the decreased level of lung pressure (estimated indirectly via the SPL measurements) in the HF patients might be a general sign of lung function reduction associated with HF. Lowering of lung function in HF can be caused by

many issues, such as the generation of edema in the lungs and constantly elevated pulmonary vascular pressures (Murphy et al., 2020). In summary, it is possible that the reported findings on the speech production characteristics in speakers with HF cannot be explained only by the generation of edema at the vocal folds (as hypothesized based on the study in Murton et al. (2017)), but they might also be due to decreased lung function caused by HF. Due to the lack of vocal fold imaging in the current study, it is, however, not possible to rule out the possibility that the reported changes in the glottal flow pulses in the HF patients might have been due to edema in the vocal folds.

The parameter distributions for the female speakers are shown in Figs. 4 and 5. From the box plots generated for the female speakers

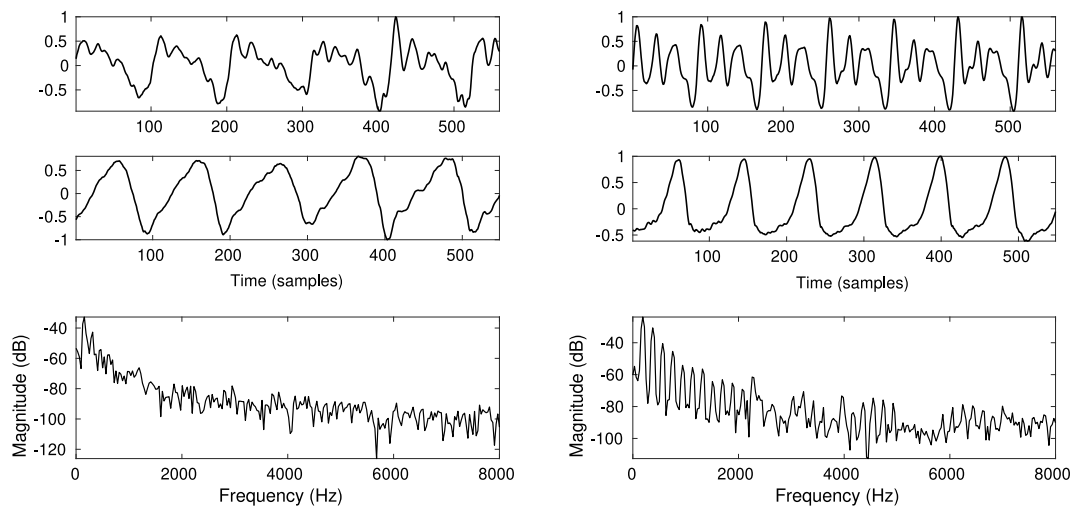


Fig. 6. An illustration of glottal flow waveforms and their spectra estimated from a vowel produced by an HF patient (left column) and from a vowel produced by a healthy speaker (right column). The three rows from top to bottom show, respectively, the time-domain speech signal, the estimated time-domain glottal flow, and the magnitude spectrum of the estimated glottal flow.

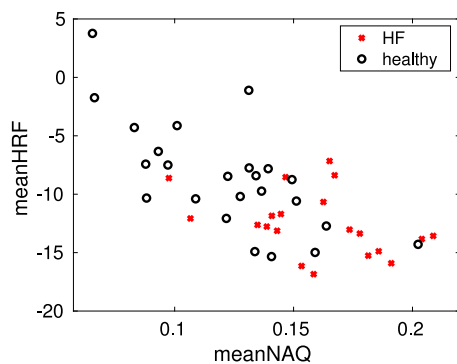


Fig. 7. Scatter plot between the mean of NAQ and the mean of HRF for the speakers with HF (red crosses) and for the healthy speakers (black circles). The mean was computed by averaging the parameter over the three segments of the vowel [a]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Figs. 4 and 5), it can be seen that the plots of most of the glottal parameters and the SPL exhibit trends similar to those in the male speakers (Fig. 2 and 3). This suggests that the characteristics of the glottal flow seem to distinguish the speakers with HF from healthy speakers irrespective of sex. Unlike for the male speakers, however, the Wilcoxon test of the female subjects indicated no statistically significant differences between the HF patients and the healthy controls for most of the glottal parameters. This discrepancy between the male and female speakers is most likely due to the small number of female talkers in the current study. Hence, the results of the statistical tests of the female speakers should be considered approximate and tentative.

In summary, the results of the statistical tests, along with the parameters' box plots, suggest that speakers with HF excite their vocal tract using a more rounded, low-frequency rich glottal flow pulse compared to healthy speakers, resulting in a weak and breathy voice. This behavior of glottal pulse is further demonstrated in Figs. 6 and 7. Fig. 6 shows examples of glottal flow waveforms and their spectra estimated using the QCP method from speech signals produced by a speaker with HF and by a healthy speaker. The figure demonstrates that the glottal flow of the healthy speaker shows a clear closed phase and a short open phase. However, for the speaker with HF, the closed phase is practically absent from the glottal flow and the shape of the glottal pulse is more symmetric and rounded. Fig. 7 shows a scatter plot for a pair of

glottal parameters obtained by averaging the parameter values over the three vowel segments for each speaker. The plot in Fig. 7 shows that the speakers with HF are located mainly in the right lower corner of the 2-dimensional glottal parameter space, which corresponds to using smooth glottal pulses with large spectral tilt. The healthy controls are located in the middle/left upper part of the space, indicating the use of glottal pulses with shorter closing phases and smaller spectral tilt. In addition, the parameters show natural deviation between individual speakers as reported in many previous studies in parameterization of the glottal flow (Holmberg et al., 1988). However, it can be seen that the deviation of the two glottal parameters between the healthy speakers is larger than the corresponding parameter deviation between the HF patients. The smaller parameter deviation in the HF patients is explained by the fact that none of these speakers were capable of producing speech with a small NAQ value combined with an HRF value close to 0 dB (i.e. the region in the left upper corner of Fig. 7).

4. Conclusion

Voice recording is a non-invasive tool that could be readily available and used in the general population to detect alterations in speech that would imply further diagnostic investigations. An increasing number of studies have been published in the past decade to investigate how voice signals produced by speakers suffering from a known disorder (e.g. Parkinson's disease, Alzheimer's disease, dysphonia, specific language impairment) differ from speech produced by healthy speakers (Kiran Reddy et al., 2020; Rusz et al., 2013; Orozco-Arroyave et al., 2016; Tu et al., 2016; Lansford et al., 2016; Lopez-de Ipiná et al., 2018). Particularly in the last couple of years, a large part of these studies have investigated machine learning (ML) methods to enable the automatic detection or severity assessment of the underlying disorder. Compared to neurodegenerative diseases, such as Parkinson's disease and Alzheimer's disease, there are, however, a lot fewer studies focusing on speech characteristics in HF patients, let alone studies investigating automatic methods to detect HF from speech. This is rather surprising since HF is known to be a worldwide health problem with an increasing prevalence (Savarese and Lund, 2017). To the best of our knowledge, the acoustic characteristics of speech signals produced by HF patients have been studied only in three previous studies (Maor et al., 2016, 2020; Murton et al., 2017) and the automatic detection of HF only in one study (Kiran Reddy et al., 2021). Despite the fact that these (few) previous studies suggest that the generation of edema (a hallmark of HF) affect voice characteristics, there are, however, no

studies investigating the effect of HF on the main acoustic excitation produced by the vocal folds, the glottal flow.

Given this situation, the current study was launched as the first investigation to focus on the characteristics of the glottal excitation in production of speech by HF patients. In particular, the goal of the current study was to study whether the glottal flow signals, estimated non-invasively from speech microphone signals using GIF, differ between HF patients and their healthy controls. Based on [Murton et al. \(2017\)](#), our hypothesis was that the generation of edema in the vocal folds and lungs of HF patients should also affect the acoustical excitation signal generated by the vocal folds of these individuals and that this phenomenon should manifest itself as a difference between the glottal flow parameters extracted from HF patients and their healthy controls. To study this hypothesis, a database consisting of speech by HF patients and by healthy talkers was collected. The speech data were parameterized using glottal parameters and SPL, and statistical analyses were conducted to investigate the behavior of the parameters between the two speaker groups. Most of the glottal parameters and SPL were found to show significant differences between the two speaker groups for male speakers. For female speakers, the same observation could be made but the small number of test subjects made the analyses approximate. Furthermore, the analyses showed that the production of speech in HF was associated with incomplete glottal closure (reflected by the increased roundness of the shape of the glottal pulse) and reduced subglottal pressure (indicated by the lowering of SPL), resulting in soft and breathy speech. These findings were as hypothesized because increased roundness of the glottal pulse shape might have been caused by the generation of edema in the vocal folds. However, based on the indirectly observed lowering of lung pressure in the HF patients, the study suggests that HF may affect speech production not only by generating edema in the vocal folds but also by lowering the lung function.

This is the first, yet preliminary study to understand how HF affects the acoustical characteristics of human speech production at the level of glottal flow and the SPL of the produced speech pressure signal. Further studies should be made in (considerably) larger patient populations in order to make sensitivity analyses for different degrees of left ventricular function (ejection fraction). The study shows interesting results that accentuate the suitability of glottal source analysis as a diagnostic tool for detection of HF, or more specifically as a screening tool for the “normal” population. Hence, the future focus may be on utilizing the glottal parameters together with speech features extracted at higher levels (such as supra-segmental features) to train effective ML algorithms for the classification of speakers with HF from healthy speakers. Additionally, the significance of glottal parameters in predicting the severity level of HF from speech signals should be studied due to its feasibility in monitoring progression of HF. This preliminary study focused on studying manually selected segments of the sustained vowel [a], which is a phone for which GIF analysis is known to be most accurate. It is, however, possible to conduct GIF analysis for any voiced segment of speech by using automatic inverse filtering and parameterization as shown, for example, by studies in speech synthesis ([Airaksinen et al., 2018](#); [Cui et al., 2018](#)). Therefore, there are no methodological obstacles to take advantage of voice source information in the similar manner as done in this study to extend the analysis to cover other voiced phones than just the vowel [a]. The main drawback of the proposed approach is that it is efficient for clinical environments but not necessarily for real-world environments. Currently, there is an increasing demand for remote and real-time monitoring of patients with HF. In realistic scenarios, however, the accuracy of GIF methods might deteriorate since the input speech might be corrupted, for example, by environmental noise or by band-pass filtering and speech compression algorithms used in speech transmission. Hence, there is a need for robust GIF methods which can yield reliable glottal flow estimates even from degraded speech. Finally, as with several other medical applications, it is worth emphasizing that

the HF detection systems based on processing speech signals need to be highly accurate. In order to improve the overall detection accuracy, speech-based systems could be combined with data collected from other sources (e.g. x-ray imaging).

CRediT authorship contribution statement

Mittapalle Kiran Reddy: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft. **Hilla Pohjalainen:** Data curation, Visualization, Writing – review & editing. **Pyry Helkkula:** Validation, Visualization, Writing – review & editing. **Kasimir Kaitue:** Validation, Visualization, Writing – review & editing. **Mikko Minkkinen:** Validation, Visualization, Writing – review & editing. **Heli Tolppanen:** Validation, Visualization, Writing – review & editing. **Tuomo Nieminen:** Validation, Visualization, Writing – review & editing. **Paavo Alku:** Conceptualization, Validation, Visualization, Supervision, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research has been funded by the Academy of Finland (project no. 330139).

References

- Airaksinen, M., Juvela, L., Bollepalli, B., Yamagishi, J., Alku, P., 2018. A comparison between STRAIGHT, glottal, and sinusoidal vocoding in statistical parametric speech synthesis. *IEEE/ACM Trans. Audio Speech Lang. Process.* 20 (9), 1658–1670.
- Airaksinen, M., Raitio, T., Story, B., Alku, P., 2013. Quasi closed phase glottal inverse filtering analysis with weighted linear prediction. *IEEE/ACM Trans. Audio Speech Lang. Process.* 22 (3), 596–607.
- Airas, M., Pulakka, H., Bäckström, T., Alku, P., A toolkit for voice inverse filtering and parametrisation. In: *Proceedings of the 9th European Conference on Speech Communication and Technology*, Lisbon, Portugal. pp. 2145–2148.
- Alku, P., 2011. Glottal inverse filtering analysis of human voice production - a review of estimation and parameterization methods of the glottal excitation and their applications. *Sadhana* 36 (5), 623–650.
- Alku, P., Airas, M., Björkner, E., Sundberg, J., 2006a. An amplitude quotient based method to analyze changes in the shape of the glottal pulse in the regulation of vocal intensity. *J. Acoust. Soc. Am.* 120 (1), 1052–1062.
- Alku, P., Bäckström, T., Vilkmann, E., 2002. Normalized amplitude quotient for parameterization of the glottal flow. *J. Acoust. Soc. Am.* 112 (2), 701–710.
- Alku, P., Pohjalainen, J., Vainio, M., Laukkanen, A.M., Story, B.H., 2013. Formant frequency estimation of high-pitched vowels using weighted linear prediction. *J. Acoust. Soc. Am.* 134 (2), 1295–1313.
- Alku, P., Story, B., Airas, M., 2006b. Estimation of the voice source from speech pressure signals: Evaluation of an inverse filtering technique using physical modelling of voice production. *Folia Phoniatr. Logop.* 58 (2), 102–113.
- Alku, P., Strik, H., Vilkmann, E., 1997. Parabolic spectral parameter - a new method for quantification of the glottal flow. *Speech Comm.* 22, 67–79.
- Alku, P., Vilkmann, E., 1996. Amplitude domain quotient for characterization of the glottal volume velocity waveform estimated by inverse filtering. *Speech Comm.* 18 (2), 131–138.
- Alvarez, R.M.B.D., Barón-López, F.J., Alguacil, M.D., Dawid-Milner, M.S., 2013. Interactions between voice fundamental frequency and cardiovascular parameters. Preliminary results and physiological mechanisms. *Logop. Phoniatr. Vocology* 38 (2), 52–58.
- Childers, D.G., Lee, C.K., 1991. Vocal quality factors: analysis, synthesis, and perception. *J. Acoust. Soc. Am.* 90 (5), 2394–2410.
- Coronel, R., Groot, J.R.D., Lieshout, J.J.V., 2001. Defining heart failure. *Cardiovasc. Res.* 50 (3), 419–422.
- Cui, Y., Wang, X., He, L., Soong, F.K., 2018. A new glottal neural vocoder for speech synthesis. In: *Proceedings of INTERSPEECH*, Hyderabad, India. pp. 2017–2021.
- Drugman, T., Alku, P., Alwan, A., Yegnanarayana, B., 2014. Glottal source processing: From analysis to applications. *Comput. Speech Lang.* 28 (5), 1117–1138.
- Espinoza, V.M., Matás, J.H., Mehta, D.D., Hillman, R.E., 2017. Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *J. Speech Lang. Res.* 60, 2159–2169.

- Fant, G., 1970. Acoustic Theory of Speech Production. Mouton, The Hague, Netherlands.
- Gauffin, J., Sundberg, J., 1989. Spectral correlates of glottal voice source waveform characteristics. *J. Speech Lang. Hear. Res.* 32 (3), 556–565.
- Gillespie, S., Logan, Y.Y., Moore, E., Laures-Gore, J., Russell, S., Patel, R., 2017. Cross-database models for the classification of dysarthria presence. In: *Proceedings of INTERSPEECH*, Stockholm, Sweden. pp. 3127–3131.
- Hecht, H.H., 1956. Heart failure and lung disease. *Circulation* 14 (2), 265–290.
- Hollander, M., Wolfe, D.A., Chicken, E. (Eds.), 2013. *Nonparametric Statistical Methods*, Vol. 751. John Wiley & Sons.
- Holmberg, E., Hillman, R., Perkell, J., 1988. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J. Acoust. Soc. Am.* 84 (2), 511–529.
- Lopez-de Ipina, K., Martinez-de Lizarduy, U., Calvo, P., Mekyska, J., Beitia, B., Barroso, N., Estanga, A., Tainta, M., Ecay-Torres, M., 2018. Advances on automatic speech analysis for early detection of alzheimer's disease: a non-linear multi-task approach. *Curr. Alzheimer Res.* 15 (2), 139–148.
- Keerthana, Y.M., Reddy, M.K., Rao, K.S., 2019. Cwt-based approach for epoch extraction from telephone quality speech. *IEEE Signal Process. Lett.* 26 (8), 1107–1111.
- Kiran Reddy, M., Alku, P., Sreenivasa Rao, K., 2020. Detection of specific language impairment in children using glottal source features. *IEEE Access* 8, 15273–15279.
- Kiran Reddy, M., Helkkula, P., Keerthana, Y.M., Kaitue, K., Minkinen, M., Tolppanen, H., Nieminen, T., Alku, P., 2021. The automatic detection of heart failure using speech signals. *Comput. Speech Lang.* 69, 101205.
- Lansford, K.L., Berisha, V., Utianski, R.L., 2016. Modeling listener perception of speaker similarity in dysarthria. *J. Acoust. Soc. Am.* 136 (6), EL209–EL215.
- Maor, E., Perry, D., Mevorach, D., Taiblum, N., Luz, Y., Mazin, I., Lerman, A., Koren, G., Shalev, V., 2020. Vocal biomarker is associated with hospitalization and mortality among heart failure patients. *J. Acoust. Soc. Am.* 9 (7), e13359.
- Maor, D.S., Lerman, L.O., Lerman, A., 2016. The sound of atherosclerosis: voice signal characteristics are independently associated with coronary artery disease. *Circulation* 134 (1), A15840.
- Monsen, R., Engebretson, A., 1977. Study of variations in the male and female glottal wave. *J. Acoust. Soc. Am.* 62 (4), 981–993.
- Morgan, C.J., 2017. Use of proper statistical techniques for research studies with small samples. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* 313 (5), L873–L877.
- Murphy, S.P., Ibrahim, N.E., Januzzi, J.L., 2020. Heart failure with reduced ejection fraction: a review. *Jama* 324 (5), 488–504.
- Murton, O.M., Hillman, R., Mehta, D., Semigran, M., Daher, M., Cunningham, T., Verkouk, K., Tabtabai, S., Steiner, J., Dec, G.W., Ausiello, D., 2017. Acoustic speech analysis of patients with decompensated heart failure: A pilot study. *J. Acoust. Soc. Am.* 142 (4), EL401–EL407.
- Narendra, N.P., Schuller, B., Alku, P., 2021. The detection of parkinson's disease from speech using voice source information. *IEEE/ACM Trans. Audio Speech Lang. Process.* 29, 1925–1936.
- Novotný, M., Pet, D., Daly, I., Evžen, R., Ruzs, J., 2020. Glottal source analysis of voice deficits in newly diagnosed drug-naïve patients with parkinson's disease: Correlation between acoustic speech characteristics and non-speech motor performance. *Biomed. Signal Process. Control* 57, 101818.
- Orlikoff, R.F., Baken, R.J., 1989. The effect of the heartbeat on vocal fundamental frequency perturbation. *J. Speech Lang. Hear. Res.* 32 (3), 576–582.
- Orozco-Arroyave, J.R., Hönig, F., Arias-Londoño, J.D., Vargas-Bonilla, J.F., Daqrouq, K., Skodda, S., Ruzs, J., Nöth, E., 2016. Automatic detection of parkinson's disease in running speech spoken in three different languages. *J. Acoust. Soc. Am.* 139 (1), 481–500.
- Ponikowski, P., Voors, A.A., Anker, S.D., Bueno, H., Cleland, J., Coats, A.J., Falk, V., Gonzalez-Juanatey, J.R., Harjola, V., Jankowska, E.A., Jessup, M., 2016. Esc guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur. Heart J.* 37 (27), 2129–2200.
- Ruzs, J., Cmejla, R., Tykalova, T., Ruzickova, H., Klempir, J., Majerova, V., Picmausova, J., Roth, J., Ruzicka, E., 2013. Imprecise vowel articulation as a potential early marker of parkinson's disease: Effect of speaking task. *J. Acoust. Soc. Am.* 134 (3), 2171–2181.
- Sapienza, C.M., Stathopoulos, E.T., Dromey, C., 1998. Approximations of open quotient and speed quotient from glottal airflow and EGG waveforms: effects of measurement criteria and sound pressure level. *J. Voice* 12 (1), 31–43.
- Sara, J.D.S., Maor, E., Borlaug, B., Lewis, B.R., Orbelo, D., Lerman, L.O., Lerman, A., 2020. Non-invasive vocal biomarker is associated with pulmonary hypertension. *PLoS One* 15 (4), e0231441.
- Savarese, G., Lund, L.H., 2017. Global public health burden of heart failure. *Cardiac Fail Rev.* 3 (1), 7–11.
- Sundberg, J., Titze, I., Scherer, R., 1993. Phonatory control in male singing: A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. *J. Voice* 7 (1), 15–29.
- Švec, J.G., Granqvist, S., 2018. Tutorial and guidelines on measurement of sound pressure level in voice and speech. *J. Speech Lang. Hear. Res.* 61 (3), 441–461.
- Thompson, L., Emily, L., Robb, M.P. (Eds.), 2011. *The impact of breathiness on the intelligibility of speech*. Poster Presentation.
- Titze, I., 1992. Phonation threshold pressure: A missing link in glottal aerodynamics. *J. Acoust. Soc. Am.* 91 (5), 2926–2935.
- Tu, M., Wisler, A., Berisha, V., Liss, J.M., 2016. The relationship between perceptual disturbances in dysarthric speech and automatic speech recognition performance. *J. Acoust. Soc. Am.* 140 (5), EL416–EL422.
- Vasan, R., Xanthakis, V., Lyass, A., Andersson, C., Tsao, C., Cheng, S., Aragam, J., Benjamin, E., Larson, M., 2018. Epidemiology of left ventricular systolic dysfunction and heart failure in the framingham study: an echocardiographic study over 3 decades. *N. Engl. J. Med.* 11 (1), 1–11.
- Verdolini, K., Min, Y., Titze, I.R., Lemke, J., Brown, K., Van Mersbergen, M., Jiang, J., Fisher, K., 2002. Biological mechanisms underlying voice changes due to dehydration. *J. Speech Lang. Hear. Res.* 45 (2), 268–281.
- Vilkman, E., Alku, P., Vintturi, J., 2002. Dynamic extremes of voice in the light of time domain parameters extracted from the amplitude features of glottal flow and its derivative. *Folia Phoniatr. Logop.* 54 (3), 144–157.
- William, C.L., Diefenbach, M.D., 1949. Left vocal-cord paralysis associated with hypertensive heart disease. *New England Journal of Medicine* 240 (11), 419–420.
- Wong, D., Markel, J., Gray, A., 1979. Least squares glottal inverse filtering from the acoustic speech waveform. *IEEE/ACM Trans. Audio Speech Lang. Process.* 27 (4), 350–355.