RESEARCH **Open Access**

Simulation of tremulous voices using a biomechanical model

Rubén Fraile[1*], Juan Ignacio Godino-Llorente[1] and Malte Kob[2]

Abstract

Vocal tremor has been simulated using a high-dimensional discrete vocal fold model. Specifically, respiratory, phonatory, and articulatory tremors have been modeled as instabilities in six parameters of the model. Reported results are consistent with previous knowledge in that respiratory tremor mainly causes amplitude modulation of the voice signal while laryngeal tremor causes both amplitude and frequency modulation. In turn, articulatory tremor is commonly assumed to produce only amplitude modulations but the simulation results indicate that it also produces a high-frequency modulation of the output signal. Furthermore, articulatory tremor affects the frequency response of the vocal tract and it might thus be detected by analyzing the spectral envelope of the acoustic signal.

Keywords: Vocal tremor; Biomechanical modeling; Voice production modeling and simulation

Introduction

Tremor can be defined as an *involuntary oscillatory movement of a body part* [1]. Correspondingly, vocal tremor can be defined as a *low-frequency fluctuation* (i.e. modulation) *in the amplitude or frequency (or both)* of voice [2]. Vocal tremor happens at a rate well below the fundamental frequency, between 1 and 15 Hz [2]. It differs from voice amplitude and frequency perturbations in that while these are random and fast deviations from stable underlying values of amplitude and fundamental frequency, vocal tremor involves an instability in such values [2]. Similarly, vibrato is a modulation of the acoustic voice signal too, but vocal tremor differs from it in that for vibrato, the span of typical modulation rates is narrower (between 4 and 7 Hz) [2] and the modulation is more cyclic [3]. While vocal tremor happens during speaking, vibrato is a feature of the singing voice [2] that may express emotions such as anger, happiness, fear, or sadness [4].

The presence of vocal tremor in the acoustic signal is better perceived in the phonation of sustained sounds than in running speech [5,6]. This is coherent with the nature of voice tremor, that is, instabilities in the voice production system are better noticed when a stable phonation task is performed. Acoustic analysis of the voice aimed at detecting, measuring, or modeling instabilities due to vocal tremor is usually approached by seeking for modulations in the sequence of pitch periods (e.g. [7]). The presence of such modulations by itself is not an indicator of dysphonia; instead, the presence of tremor is necessary for a natural sounding voice [2]. In fact, vocal tremor in healthy individuals (physiological tremor) has been measured, showing modulation rates below 5 Hz [8]. However, the modulation rate is not definite in discriminating between physiological and pathological voice tremors. Jiang et al. and Anand et al. showed that for pathological voice tremors, with rates between 4 and 7 Hz, the difference with physiological voice tremor was more in the modulation extent (or modulation amplitude) than in the modulation rate [9,10].

The fact that vocal tremor is more related to modulation extents than to modulation rates is consistent with results indicating that the severity of tremor is correlated with acoustic measures of amplitude and frequency perturbations [11]. This is due to perturbation measures such as *jitter* or *shimmer* not including information on the autocorrelation properties of the varying magnitude; i.e. they are perturbation functions of order 0 [12]. As a consequence, they measure the same the extent of random and rapid perturbations as the extent of lower frequency modulations.

*Correspondence: rfraile@ics.upm.es
[1] Signal Theory & Communications Department, Universidad Politécnica de Madrid, Campus Sur - Carretera de Valencia km 7, 28031 Madrid, Spain
Full list of author information is available at the end of the article

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 2 of 12

Regarding perception, frequency modulation seems to be the most relevant feature for the perceptual detection of vocal tremor [10,13] but tremor is only perceived if its intensity (i.e. modulation extent) is above a certain threshold [6].

From the point of view of the anatomy of the voice production system, vocal tremor may be produced by three different sources [14,15]:

- *Respiratory system:* Instabilities in the respiratory system can typically be detected by analyzing the voice intensity contours [16]. Some results from the computer simulation of respiratory vocal tremor also point out that this is more related to amplitude modulations than to frequency modulations [15], but more recent measurements from *in vivo* induction of respiratory tremor indicate that it also affects fundamental frequency [17].

- *Phonatory system:* To the greatest extent, tremor in the phonatory system, or laryngeal tremor, is caused by irregular tension patterns in laryngeal muscles [18], mainly the thyroarytenoid (TA), the cricothyroid (CT), and the inter-arytenoid (IA) (cf. Figure 1). The activity of the laryngeal muscles has a relevant impact on the fundamental frequency of voice [19-21], although the effect of interactions between such muscles is not fully understood yet [22] and the relation between muscles tension and

fundamental frequency is complex [23]. Accordingly, the main effect of laryngeal tremor is commonly assumed to be the frequency modulation of the acoustic wave [24]. However, amplitude modulation is also expected to happen due to the interaction between harmonics and formants [25].

- *Articulatory system:* Articulatory tremor is produced by instabilities in the position of the articulators (e.g. jaw or tongue), position of the epiglottis, width of the epilarynx, etc. Similarly to respiratory tremor, articulatory tremor is thought to affect voice amplitude [24] but also to produce correlations between airflow magnitude and pressure wave amplitudes different to the case of respiratory tremor [14].

Regarding simulation of vocal tremor, only a limited number of experiments have been reported in the scientific literature up to now. Hanquinet et al. developed a synthesizer of disordered voices in which simulation of vocal tremor was possible [26]. Their simulation approach was based on modeling the acoustic voice signal, not the underlying biomechanical process that produces it. Zhang and Jiang used a two-mass model of the vocal folds to produce simulated tremulous voices with the objective of studying the nonlinear dynamics of vocal tremor [27]. Lester, Barkmeier-Kraemer, and Story utilized a kinematic model of the vocal folds to simulate tremor produced
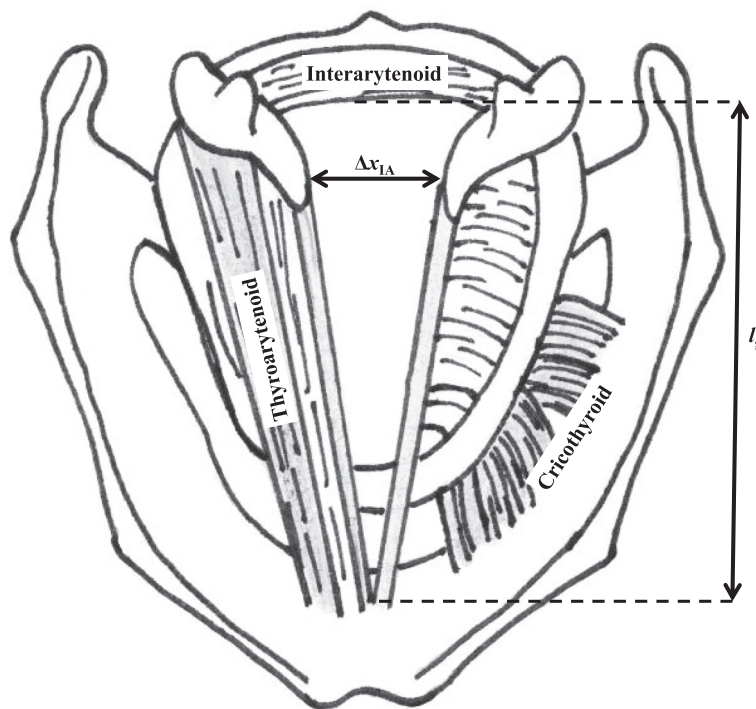
Figure 1 Larynx diagram. Sketch of the larynx with indication of intrinsic laryngeal muscles TA, CT, and IA.

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 3 of 12

by different anatomic sources [15,24]. Using a different approach, Jiang et al. [14] and Lester and Story [17] made use of *in vivo* induction of tremor to seek for relationships between acoustic features and tremor sources.

This paper reports on the results of using a high-dimensional multiple-mass vocal fold model for simulating vocal tremor. In comparison to other options, using this model allows to study the impact that the addition of tremor to specific biomechanical parameters has on the acoustic voice signal. Namely, the paper includes results from the simulation of tremulous behavior in lung pressure, vocal fold stiffness, vocal ligament stress, vocal fold length, vocal fold adduction, and jaw opening. Both respiratory [15] and phonatory tremors [24,27] had been simulated before but, to the best of authors⬛ knowledge, this paper is the first one that reports results on the simulation of articulatory tremor. Results are presented in terms of modulation extents, both for amplitude and for frequency modulations, bearing in mind the idea that diverse anatomic sources of tremor should produce different modulation extents [24]. In the case of articulatory tremor, it is shown that amplitude and frequency modulations are not the only acoustic effects of tremor; additionally, articulatory tremor generates instabilities in the spectral envelope of the voice signal.

Voice production model

A detailed and complete description of the voice production model can be found in [28]. Here, only a brief and qualitative description is included, aimed at providing a framework for a better understanding of the tremor model. This voice production model belongs to the class of discrete, high-dimensional, biomechanical models. It is based on previous works carried out by Kob et al. [29,30], and it has been successfully employed for analyzing the pathogenesis of vocal fold nodules [31] and for simulating laryngeal disorders [28]; a resembling model was also used by Wong et al. [32] for a similar purpose.

In the biomechanical model used, each vocal fold is formed by a set of 15 contiguous elements aligned in the anterior-posterior direction. Thus, each element models the transverse section of the vocal fold in a different coronal plane. In turn, each element is composed by two masses: the upper mass that approximately models the mucosa and the lower mass that somewhat models the vocal muscle (TA muscle) and the vocal ligament. Simulated mechanical properties of the vocal-fold tissues include tissue elasticity, elasticity of the interfaces between adjacent tissues, and compression forces tending to recover form after collision. Other simulated forces are the vocal ligament stress, the pressure induced by the glottal airflow, and the subglottal pressure.

The model does not include simulation of the subglottal pressure waves and resonances, so subglottal pressure

is made equal to the lung pressure. As for supraglottal structures, the vocal tract is modeled as a set of 44 concatenated cylinders having different cross-section areas. Pressure wave propagation along the vocal tract has been simulated using the Kelly-Lochbaum model plus an energy loss factor applied at each cylinder. Wave radiation at the output of the last cylinder has been simulated, assuming that no external acoustic wave arrives to the lips.

All parameters of the voice production model have been assigned the values specified in [28], except for those summarized in Table 1. These changes have been made with the purpose of increasing the naturalness on the synthesized sound waves.

Tremor model

Respiratory tremor

As mentioned before, vocal tremor may have three different anatomic sources: the respiratory system, the phonatory system, and the articulatory system. Respiratory tremor, which is caused by irregularities in the behavior of respiratory muscles, has been modeled by making the lung pressure (P_{sub} in [28]) variable instead of constant, that is:

$$P_{\text{sub}}(t) = P_0 + p(t) \tag{1}$$

where $P_0 = 700$ Pa as in [28]. Although the subglottal pressure is to some extent related to fundamental frequency, it is not a primary cause for changes in pitch [19,20]. Thus, *a priori* the main expected effect of the time variability of P_{sub} is a modulation in the amplitude of the acoustic signal [15,16].

Phonatory tremor

At laryngeal level, phonatory tremor is conjectured to be produced by instabilities in the tension patterns of intrinsic laryngeal muscles (e.g. TA and CT). Results reported by Finnengan et al. indicate that these are more related to vocal tremor than extrinsic laryngeal muscles [18]. The activity of the CT muscle (Figure 1) is crucial for

Table 1 Values of simulation model parameters that have been changed with respect to [28]

Description	Parameter	Value
Height of the upper masses	ΔMz	1.0 mm
Height of the lower masses	ΔVz	2.2 mm
Depth of the upper masses	ΔMx	1.8 mm
Depth of the lower masses	ΔVx	2.2 mm
Sum of the upper masses at each side	$\sum_{i=1}^{N} m_{Msi}$	25.7 mg
Sum of the lower masses at each side	$\sum_{i=1}^{N} m_{Vsi}$	70.5 mg
Stiffness of the link between boundary and lower mass	k_{sVB}	2.0 N/m
Stiffness of the link between upper and lower mass	k_{sMV}	1.5 N/m

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 4 of 12

the determination of the voice fundamental frequency [19-21]. Its function is twofold [23]: on the one hand, it affects the tension and the elongation of the vocal cords (i.e. vocal ligaments); on the other hand, it helps in modifying the stiffness of the vocal folds. This second function is shared with the TA muscle, which is parallel to the vocal ligament (Figure 1). Although TA activity exhibits a moderately high degree of correlation with CT activity, the TA muscle seems to play a secondary role in the control of fundamental frequency as compared with the CT muscle [21].

The activities of the TA and CT muscles can be modeled in discrete low-dimensional vocal fold simulation models as changes in the mass and stiffness of the vocal fold elements [23,33]. The use of a high-dimensional model allows to specifically model the changes in tension and elongation of the vocal ligament caused by the CT muscle. Hence, tremor induced by instabilities in the activity of TA and CT muscles is modeled by variations in the stiffness of the vocal folds and tension and length of the vocal ligament.

Changes in stiffness have been simulated by adding variability to the stiffness factors k_{sMV} and k_{sVB} in [28]. This variability has been modeled as a multiplicative factor:

$$k_s(t) = k_{s0} \cdot (1 + k(t)) \tag{2}$$

where k_{s0} has the values specified in Table 1 as k_{sVB} and k_{sMV}. Instabilities in the tension of the vocal ligament have been simulated by adding a variable term to the maximum active stress:

$$\sigma_{MAX}^{act}(t) = \sigma_{MAX0}^{act} + \sigma(t) \tag{3}$$

where $\sigma_{MAX0}^{act} = 60$ kPa. Last, irregularities in the elongation of the vocal cords have been modeled by multiplying the glottal length l_g (cf. Figure 1) by a variable factor, while correcting the dimensions of the vocal fold elements in order to keep their mass constant during simulations:

$$l_g(t) = l_{g0} \cdot (1 + l(t)) \tag{4}$$

with $l_{g0} = 14$ mm.

In addition to the role of TA and CT muscles in the production of vocal tremor, some research results indicate that the role of the IA muscle should not be underestimated [34]. The IA muscle, similarly to the CT, affects tension and elongation of the vocal ligaments. But it also controls the adduction of the vocal folds. Therefore, the inter-arytenoid distance, Δx_{IA} (cf. Figure 1), has been added to the variables above in order to account for the role of the IA muscle:

$$\Delta x_{IA}(t) = \Delta x(t) \tag{5}$$

where it has been assumed that the default configuration is $\Delta x_{IA} = 0$.

Articulatory tremor

Supraglottal structures, i.e. vocal tract, may contribute to the production of vocal tremor. This is hypothesized to be the case of patients suffering from Parkinson⊠s disease [9]. In fact, jaw and lip tremors are some of the features that help in differentiating between parkinsonian and essential tremor [35,36].

Vocal tract configuration, as in the case of the larynx, is multi-dimensional, that is, it depends on the values of several variables [37], such as the positions of jaw, hyoid, tongue body, tongue blade, lips, velum, etc. In the simulation model used, the vocal tract is mainly described by geometric parameters, specifically cross sections of concatenated elements. Therefore, articulatory tremor is modeled by making such geometric parameters variable.
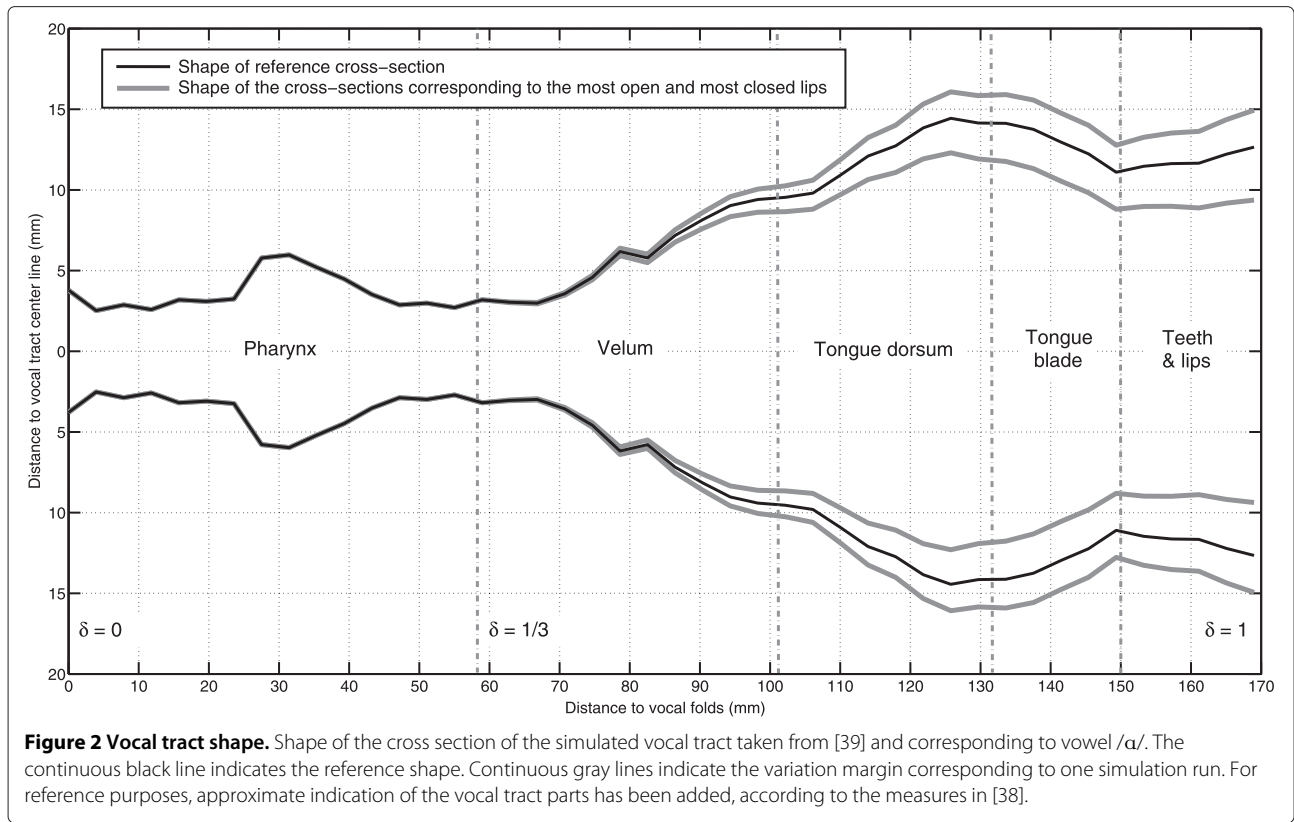
Since jaw tremor is one specific feature of parkinsonian tremor, instabilities in the jaw position have been simulated. Jaw movement is modeled in [37] as an angular displacement of the lower incisors with respect to a fix point. In a similar way, jaw tremor has been simulated as an angular displacement in the boundaries of the elements forming the upper part of the vocal tract. The pharynx section, which approximately equals 1/3 of the total vocal tract length [38], has been assumed to remain unaffected by jaw position in the simulation model. Consequently, the sections corresponding to the bottom third of the vocal tract have been kept unaltered. The remaining cross sections have been made variable according to the following rule:

$$S_\delta(t) = S_{\delta 0} \cdot \left(1 + \frac{\delta - \frac{1}{3}}{\frac{2}{3}} \cdot S(t)\right) = S_{\delta 0} \cdot \left(1 + \frac{3\delta - 1}{2} \cdot S(t)\right) \tag{6}$$

where δ indicates the position along the vocal tract axis ($\delta = \frac{1}{3}$ at the top of the pharynx; $\delta = 1$ at the lips) and $S_{\delta 0}$ are the corresponding reference cross sections provided by [39]. Figure 2 gives an overview of the reference shape of the vocal tract cross section and the variability induced by jaw tremor.

Mathematical model

While tremor is often considered to be a purely oscillatory movement [1,40], some analyses indicate that tremor may not always be oscillatory [13,41]. In this paper, tremor has been simulated as random, having a certain bandwidth Δf, instead of sinusoidal. Thus, when referring to the simulation results, the term *tremor bandwidth* will be preferred to *modulation rate*. Random values for each tremor variable described above ($p(t)$, $k(t)$, $\sigma(t)$, $l(t)$, $S(t)$, and $\Delta x(t)$) have been generated every $1/\Delta f$ seconds, according to the distributions mentioned below.

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 5 of 12

Figure 2 Vocal tract shape. Shape of the cross section of the simulated vocal tract taken from [39] and corresponding to vowel /ɑ/. The continuous black line indicates the reference shape. Continuous gray lines indicate the variation margin corresponding to one simulation run. For reference purposes, approximate indication of the vocal tract parts has been added, according to the measures in [38].

In order to calculate tremor values at intermediate time instants, a quadratic interpolation has been applied that ensures continuity of the first derivative.

Except for the case of the inter-arytenoid distance, all variable magnitudes that model tremor ($p(t)$, $k(t)$, $\sigma(t)$, $l(t)$, and $S(t)$) have been assumed to have Gaussian distributions with zero mean. In the case of the inter-arytenoid distance, it has been preferred to avoid negative values for $\Delta x(t)$ so values for it have been obtained by squaring values obtained from a Gaussian distribution, resulting in a χ^2 distribution with one degree of freedom.

Simulation parameters and acoustic analysis

The voice production and tremor models described in preceding sections have been used to generate voice signals with duration equal to 2 s and sampling rate equal to 8,000 Hz. In order to isolate the effects of different anatomical sources of tremor, time variability has only been modeled in one of the six aforementioned variables at a time. Simulations have been realized for four different bandwidths: Δf (Hz) $\in \{2.5, 5, 7.5, 10\}$ and diverse variances (i.e. modulation extents). Figure 3 shows one result of such simulations with its correspondent amplitude contour and fundamental frequency. The fundamental frequency has been estimated with a resolution of one sample, which at a sampling rate of 8,000 Hz and

with voices having a mean fundamental frequency around 150 Hz corresponds to a resolution equal to approximately 3 Hz.
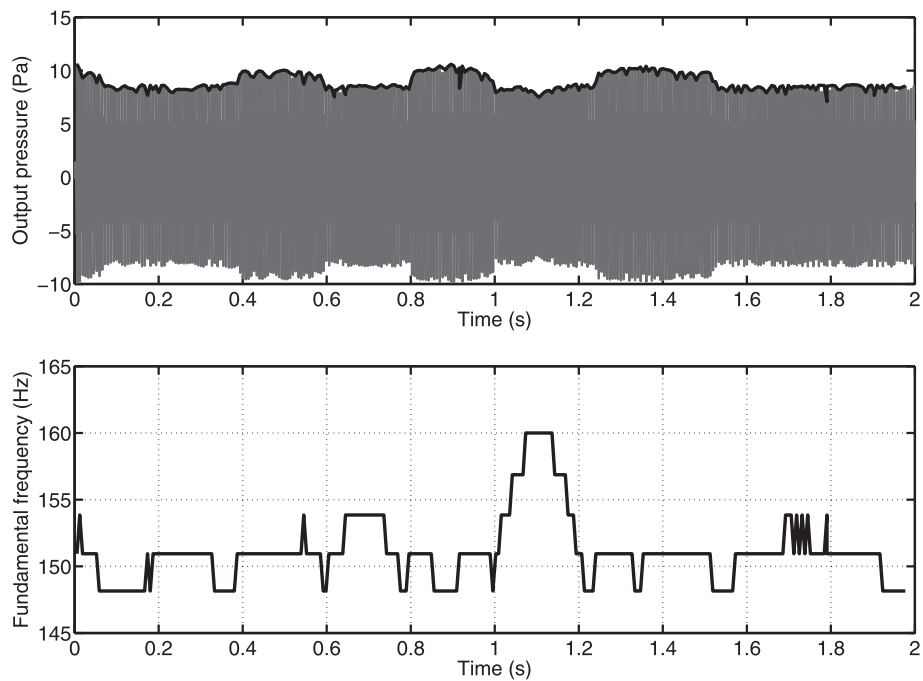
For each simulation, the sample normalized standard deviation of the corresponding variable has been computed as:

$$\overline{\sigma}_x = \frac{\sqrt{E\left\{[x(t) - E\{x(t)\}]^2\right\}}}{E\{x(t)\}} \tag{7}$$

In the case of the inter-arytenoid distance, since the mean and the standard deviation are not independent for a χ^2 distribution, no normalization has been applied to the standard deviation:

$$\sigma_{\Delta x} = \sqrt{E\left\{[\Delta x(t) - E\{\Delta x(t)\}]^2\right\}} \tag{8}$$

Periods in the output voice signal have been identified using a simple autocorrelation-based pitch detector. Based on the periods identified, two signals have been generated: one containing information on the period peak amplitude, $A(t)$, and another one containing information about the inverse period duration, $f(t)$. $A(t)$ is the amplitude modulating signal while $f(t)$ is the frequency modulating signal. Since $A(t)$ and $f(t)$ are expected to have random values

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 6 of 12

Figure 3 Simulated voice. Sample of simulated voice signals: acoustic pressure with highlighted intensity contour (top) and fundamental frequency (bottom).

obtained from unbounded distributions, the modulation extent has been defined as the normalized standard deviation of their values along a given voice signal: $\overline{\sigma}_f$ for frequency modulation and $\overline{\sigma}_A$ for amplitude modulation (Equation 7).

Results

Figure 4 shows the results of simulating respiratory tremor. For low variability in lung pressure ($\overline{\sigma}_p < 2\%$), only amplitude modulation is produced in the voice signal. For higher variabilities, frequency modulation appears but with a limited extent: $\overline{\sigma}_f < 2\%$. The extent of amplitude modulation $\overline{\sigma}_A$ varies from 3% to 20%, which means that it is from 5 to 38 times higher than the extent of frequency modulation.

Results corresponding to phonatory tremor are summarized in Figures 5, 6, 7 and 8. The impact of stiffness variations on frequency modulation is more relevant than in the case of respiratory tremor, that is, $\overline{\sigma}_A$ is only 3 to 10 times higher than $\overline{\sigma}_f$ (Figure 5). Tremor induced by instabilities in vocal fold adduction has a similar behavior (Figure 6) but with $\overline{\sigma}_f$ reaching values similar to $\overline{\sigma}_A$ in some cases. The active stress applied to the vocal ligament only has a moderate effect on vocal tremor (Figure 7), and it provides similar values for $\overline{\sigma}_f$ and $\overline{\sigma}_A$ if $\overline{\sigma}_\sigma \geq 10\%$. Last, a varying length of the vocal folds (Figure 8) seems to have a relevant effect on frequency modulation ($\overline{\sigma}_f \approx 10\%$) when the length variability goes above a

certain threshold ($\overline{\sigma}_l > 1.3\%$). Below that threshold, only amplitude modulation has a certain relevance.

Finally, the effect of articulatory tremor on the modulations of the acoustic signal is depicted in Figure 9. Similarly to the case of respiratory tremor, the effect on amplitude modulation is much more relevant than the effect on frequency modulation, with $\overline{\sigma}_A$ being at least 10 times higher than $\overline{\sigma}_f$. The limited extent of frequency modulation can be noticed in the harmonics above the fifth one in the example illustrated in Figure 10. A second effect of articulatory tremor that can be measured in the acoustic signal is the instability in the frequency response of the vocal tract. This can be noticed by smoothing the spectrum using local averaging (Figure 11).
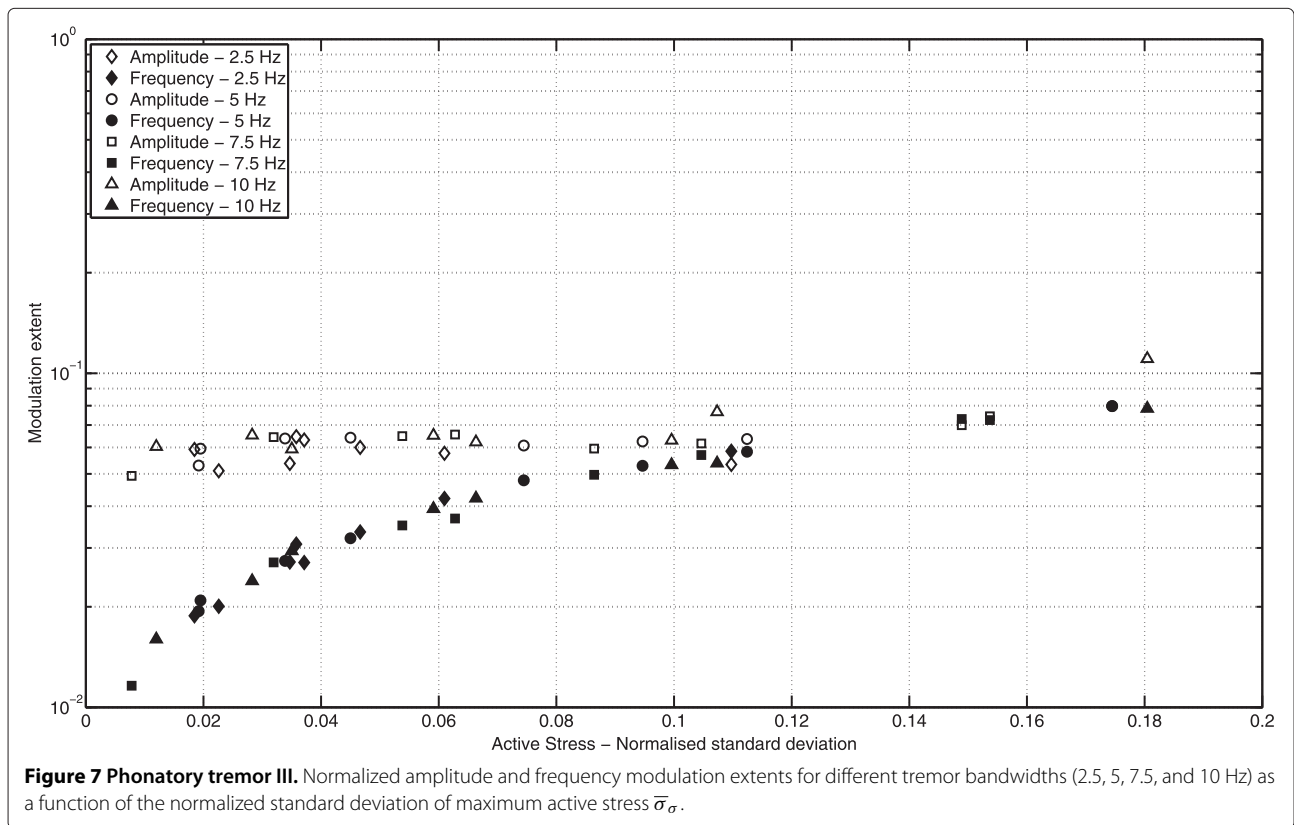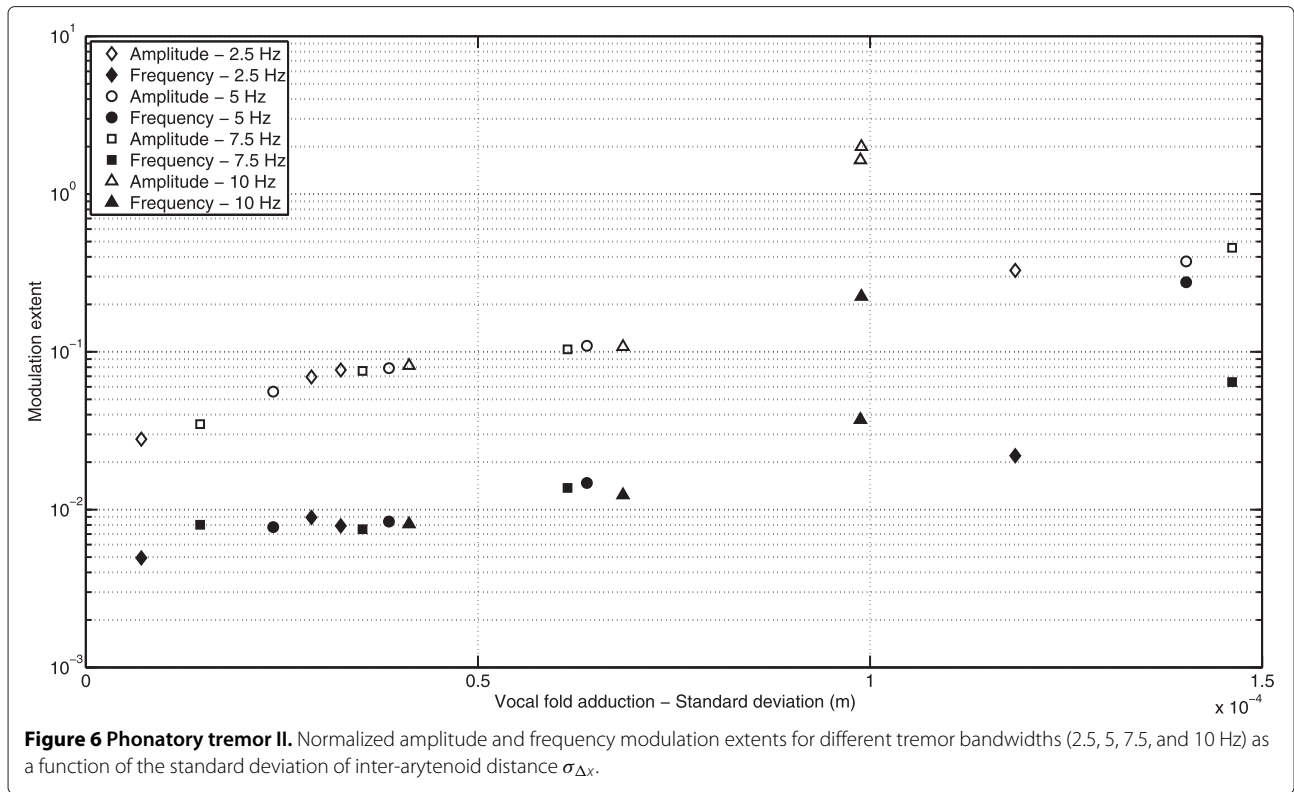
Discussion

The main effect of vocal tremor on the acoustic signal is a combination of amplitude and frequency modulation [10,13]. In contrast, other perturbation measures such as noise seem not to be related to tremor [10]. As for modulation parameters, modulation extent has been reported to be much more significant for the perception of tremor than modulation rate [9,10]. This is coherent with the fact that physiologic and pathological tremors happen within overlapped frequency ranges [1,8,9]. Thus, amplitude serves better the purpose of discriminating between them. Accordingly, the normalized standard deviations of both peak amplitude and fundamental frequency of the
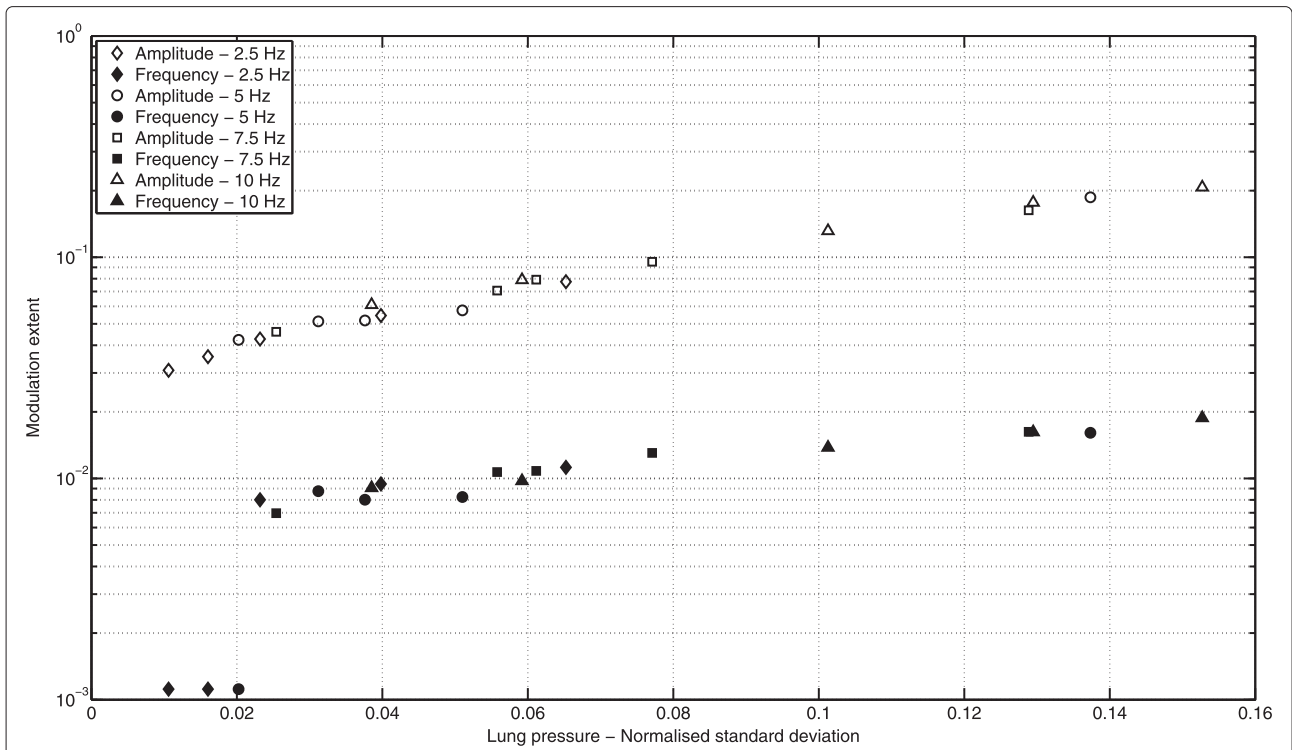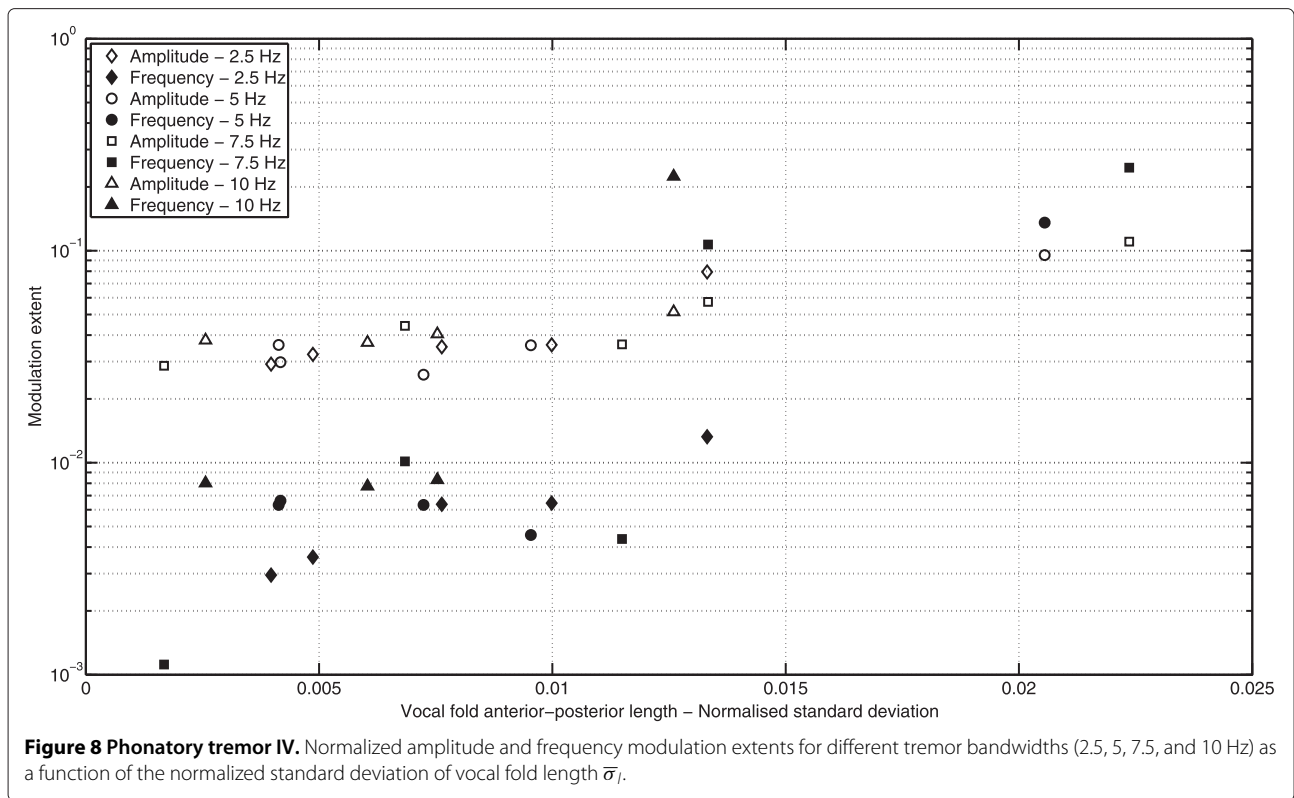
Figure 4 Respiratory tremor. Normalized amplitude and frequency modulation extents for different tremor bandwidths (2.5, 5, 7.5, and 10 Hz) as a function of the normalized standard deviation of lung pressure $\overline{\sigma}_p$.
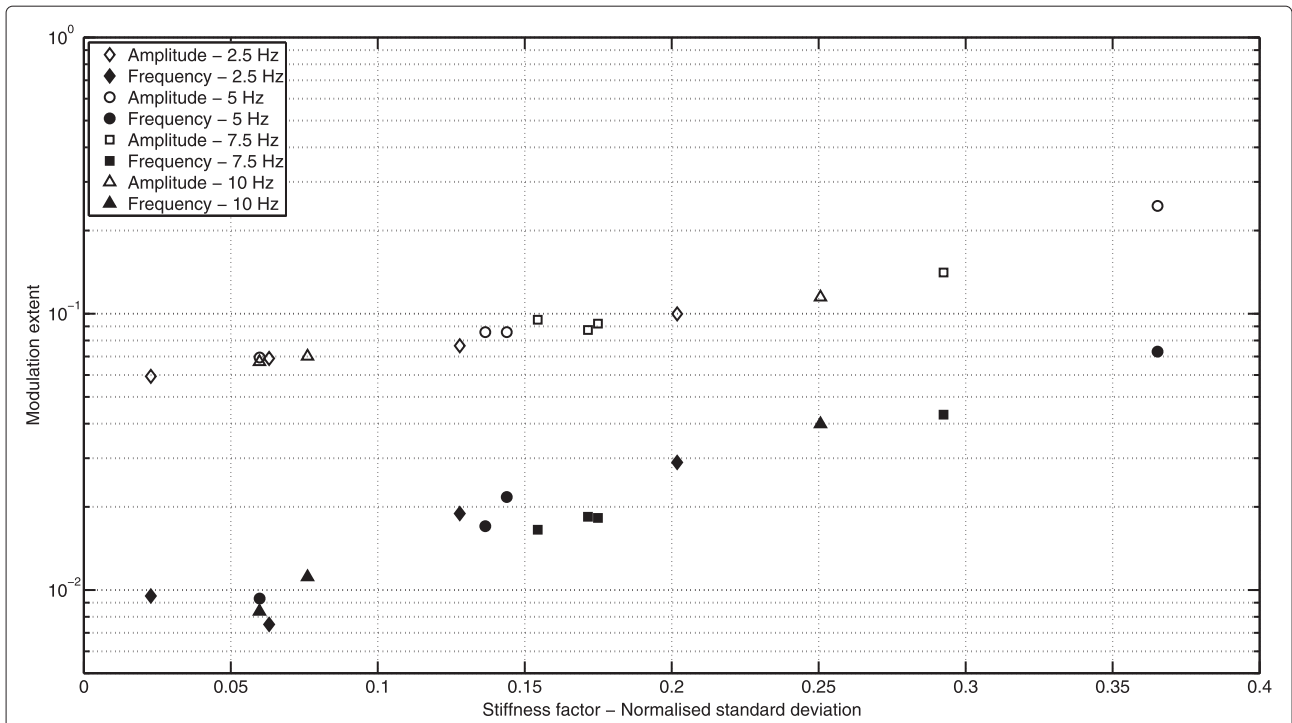
Figure 5 Phonatory tremor I. Normalized amplitude and frequency modulation extents for different tremor bandwidths (2.5, 5, 7.5, and 10 Hz) as a function of the normalized standard deviation of stiffness factor $\overline{\sigma}_k$.

Figure 6 Phonatory tremor II. Normalized amplitude and frequency modulation extents for different tremor bandwidths (2.5, 5, 7.5, and 10 Hz) as a function of the standard deviation of inter-arytenoid distance $\sigma_{\Delta x}$.

Figure 7 Phonatory tremor III. Normalized amplitude and frequency modulation extents for different tremor bandwidths (2.5, 5, 7.5, and 10 Hz) as a function of the normalized standard deviation of maximum active stress $\overline{\sigma}_{\sigma}$.

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 9 of 12

Figure 8 Phonatory tremor IV. Normalized amplitude and frequency modulation extents for different tremor bandwidths (2.5, 5, 7.5, and 10 Hz) as a function of the normalized standard deviation of vocal fold length $\overline{\sigma}_l$.
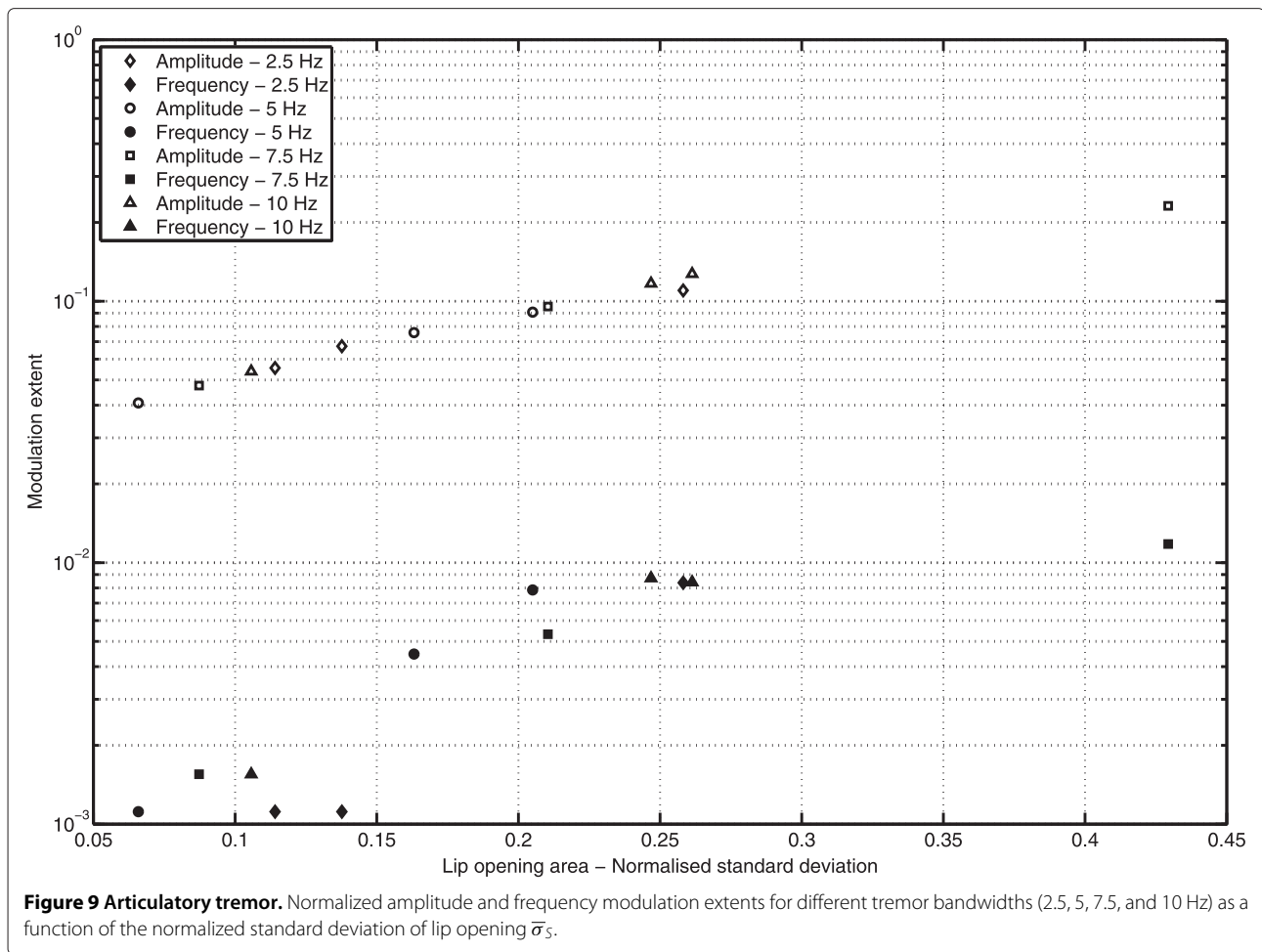
voice signal have been chosen as relevant measures of tremor for the analysis of results.

The effect of muscular tremor has been simulated independently for six physical magnitudes affecting voice production at subglottal (lung pressure), laryngeal (vocal fold stiffness, vocal fold length, vocal fold adduction, and vocal ligament stress), and supraglottal (jaw position) levels. As reported in previous studies [14,15], obtained results indicate that respiratory tremor has a relevant effect on voice amplitude but it only has a limited effect on fundamental frequency. The dominance of amplitude modulation over frequency modulation in respiratory tremor has also been recently mentioned by Lester et al. [24]. The same authors quantified the modulation extents after their simulation experiments reported in [17]: amplitude modulation extent between 5% and 50% and frequency modulation extent between 0.5% and 3.5%. Results plotted in Figure 4 are within the same range: $\overline{\sigma}_A \approx 20\%$ and $\overline{\sigma}_f \approx 2\%$ at the right of the graph.

Incidentally, the behavior of $\overline{\sigma}_f$ with respect to $\overline{\sigma}_p$ is consistent with the observation that there seems to be a threshold for the perception of respiratory tremor [6]. That is, although not formally demonstrated, it is commonly assumed that frequency modulation is more relevant for perception than amplitude modulation [13,25]. Furthermore, Kreiman et al. showed that modulation extent affects perception more than modulation rate [13].

Therefore, the jump in modulation extent that can be observed in Figure 4 near $\overline{\sigma}_p = 2\%$ is consistent with the existence of a threshold for the perception of respiratory tremor, as reported by Farinella et al. [6]. A similar threshold appears for articulatory tremor (Figure 9).

At laryngeal level, a tremulous behavior of the TA, CT, and IA muscles has a direct impact on the dynamics of the vocal folds. Such impact has been modeled by adding variability to the stiffness of the vocal folds, the stress of the vocal ligament, the length of the vocal folds, and the distance between the arytenoid cartilages on adduction. Among the four parameters, active stress in the vocal ligament seems to have the least effect on tremor (Figure 7). However, no single laryngeal parameter is likely to change independently from the rest under the joint action of laryngeal muscles. What is most relevant in the reported results is that the magnitude of frequency modulation relative to amplitude modulation is larger than for respiratory tremor, though the extents of amplitude modulations are still greater. This happens for each one of the four studied laryngeal parameters, except for the active stress, for which amplitude and frequency modulations reach similar extents when $\overline{\sigma}_\sigma \geq 10\%$. Thus, an analysis of the relative magnitudes of amplitude modulation extent and frequency modulation extent should allow to discriminate respiratory from laryngeal tremor. These results are coherent with the conclusions in [24], although
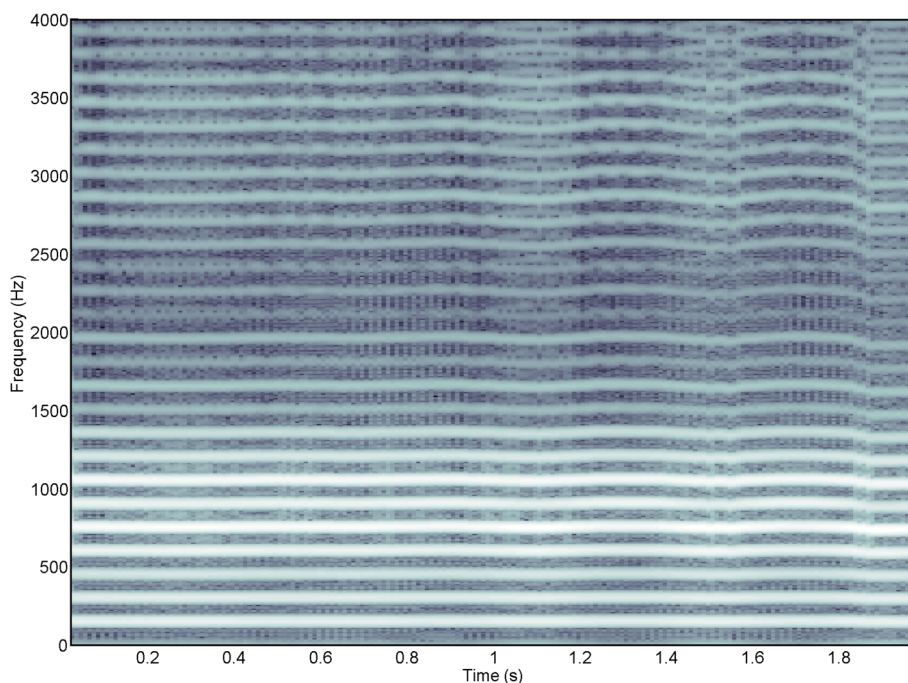
Figure 9 Articulatory tremor. Normalized amplitude and frequency modulation extents for different tremor bandwidths (2.5, 5, 7.5, and 10 Hz) as a function of the normalized standard deviation of lip opening $\overline{\sigma}_S$.

they indicate that amplitude modulations may also have a phonatory source.

Articulatory tremor may happen due to changes in shape and elastic properties at any segment of the vocal tract from the epiglottis to the lips. Among all possible models for vocal tract instability, changes in the cross section of the vocal tract segment above the pharynx have been simulated, hence modeling instabilities in the jaw position. While articulatory tremor is expected to produce only amplitude modulations of the acoustic wave [9,24], the reported results (Figure 9) indicate that frequency modulation also happens, though with a more limited extent. The classical source-filter model assumes independence between the glottal source and the vocal tract. This implies that the vocal tract cannot affect the fundamental frequency of phonation, that is, it can only amplify or attenuate signal harmonics. The expectation that articulatory tremor only causes amplitude modulations in the voice is based on this assumption. However, the existence of interaction between the glottal source and vocal tract has already been shown in several experiments [42]. As a result, vocal tract configuration can

actually affect glottal flow characteristics. In the utilized simulation model, such non-linear relation is modeled to a limited extent by (Equation 23, [28]). By including this source-filter interaction in the model, it has been shown that articulatory tremor also results in frequency modulation, although the extent of this modulation is one order of magnitude smaller than the extent of the amplitude modulation (Figure 9). On the contrary, the voice production simulator does not include models for either subglottal resonances or non-linear interaction between the trachea and larynx. As a consequence, the extent of frequency modulation due to respiratory tremor may be underestimated in the results reported here.

A second effect of articulatory tremor, illustrated in Figure 11, is the instability in the frequency response of the vocal tract. As far as the authors know, this has not been used as a measure of tremor yet. However, this effect is specific from articulatory tremor and, consequently, any measure of it could reasonably be used to discriminate whether a certain tremor comes from the articulatory structures or not.

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 11 of 12

Figure 10 Articulatory tremor - spectrogram. Spectrogram of an acoustic voice signal produced as a result of articulatory tremor simulation (tremor bandwidth equal to 7.5 Hz).

Conclusions

A high-dimensional discrete voice production model has been used to simulate vocal tremor from respiratory, phonatory, and articulatory sources. Results for respiratory and phonatory tremor are consistent with the previous knowledge of respiratory tremor mainly causing amplitude modulations in the voice signal and frequency modulation coming from phonatory tremor. However, in contrast to some previous assumptions, phonatory tremor in these results causes amplitude modulations too. Apart
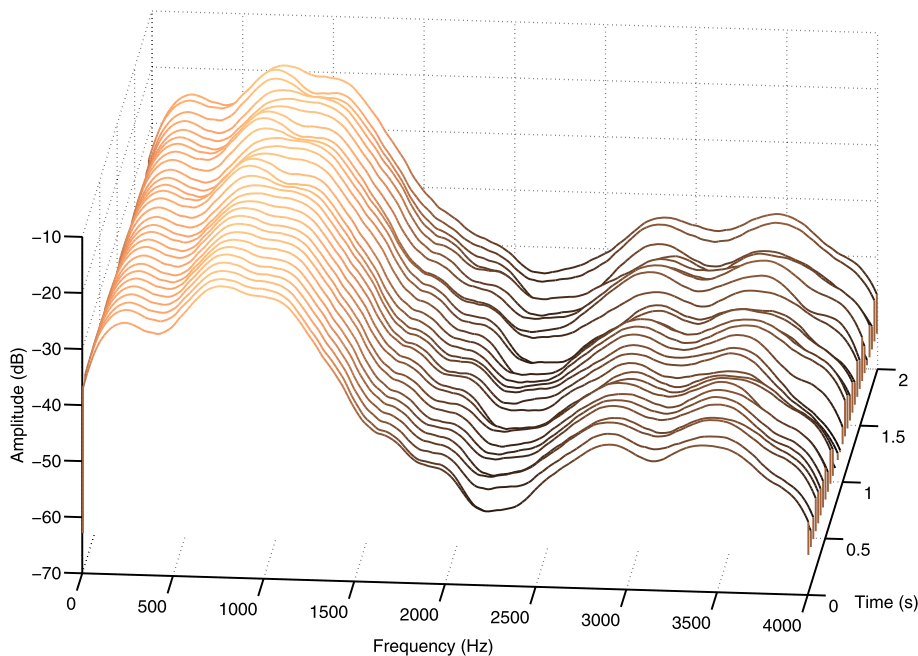
Figure 11 Articulatory tremor - smoothed spectrum. Time evolution of the smoothed spectrum of an acoustic voice signal produced as a result of articulatory tremor simulation (tremor bandwidth equal to 7.5 Hz). All plotted spectra have been normalized in energy.

Fraile *et al. EURASIP Journal on Audio, Speech, and Music Processing* (2015) 2015:1

Page 12 of 12

from the usage of a high-dimensional discrete model, the second novelty of the paper is the simulation of articulatory tremor. Similarly to the case of respiratory tremor, its main effect is amplitude modulation of the voice signal but it also causes frequency modulation with a lower modulation extent. Another specific feature of articulatory tremor that can help in differentiating it from respiratory tremor is the instability induced in the frequency response of the vocal tract.

Author details
[1]Signal Theory & Communications Department, Universidad Politécnica de Madrid, Campus Sur - Carretera de Valencia km 7, 28031 Madrid, Spain. [2]Erich-Thienhaus-Institut, Hochschule für Musik Detmold, Neustadt 22, 32756 Detmold, Germany.

References
1. G Deuschl, P Bain, M Brin, Consensus statement of the Movement Disorder Society on tremor. Mov. Disord. **13**(S3), 2–23 (1998)
2. IR Titze, Workshop on acoustic voice analysis: Summary statement. Technical report, National Center for Voice and Speech (1995)
3. C Dromey, ME Smith, Vocal tremor and vibrato in the same person: acoustic and electromyographic differences. J. Voice. **22**(5), 541–545 (2008)
4. PN Juslin, P Laukka, Communication of emotions in vocal expression and music performance: Different channels, same code? Psychol. Bull. **129**(5), 770–814 (2003)
5. A Lederle, J Barkmeier-Kraemer, E Finnegan, Perception of vocal tremor during sustained phonation compared with sentence context. J. Voice. **26**(5), 668–19 (2012)
6. KA Farinella, TJ Hixon, JD Hoit, BH Story, PA Jones, Listener perception of respiratory-induced voice tremor. Am. J. Speech Lang. Pathol. **15**(1), 72–84 (2006)
7. E Yair, I Gath, On the use of pitch power spectrum in the evaluation of vocal tremor. Proc. IEEE. **76**(9), 1166–1175 (1988)
8. J Schoentgen, Modulation frequency and modulation level owing to vocal microtremor. J. Acoust. Soc. Am. **112**(2), 690–700 (2002)
9. J Jiang, E Lin, DG Hanson, Acoustic and airflow spectral analysis of voice tremor. J. Speech Lang. Hearing Res. **43**(1), 191–204 (2000)
10. S Anand, R Shrivastav, JM Wingate, NN Chheda, An acoustic-perceptual study of vocal tremor. J. Voice. **26**(6), 811–17 (2012)
11. J Gambo, FJ Jiménez-Jiménez, A Nieto, I Cobeta, A Vegas, M Ortí-Pareja, T Gasalla, JA Molina, E García-Albea, Acoustic voice analysis in patients with essential tremor. J. Voice. **12**(4), 444–452 (1998)
12. NB Pinto, IR Titze, Unification of perturbation measures in speech signals. J. Acoust. Soc. Am. **87**(3), 1278–1289 (1990)
13. J Kreiman, B Gabelman, BR Gerratt, Perception of vocal tremor. J. Speech Lang. Hear. Res. **46**(1), 203–214 (2003)
14. J Jiang, E Lin, J Wu, C Gener, DG Hanson, Effects of simulated source of tremor on acoustic and airflow voice measures. J. Voice. **14**(1), 47–57 (2000)
15. J Barkmeier-Kraemer, BH Story, Conceptual and clinical updates on vocal tremor. The ASHA Leader. **1**, 16–19 (2010)
16. H Ackermann, W Ziegler, Cerebellar voice tremor: an acoustic analysis. J. Neurol. Neurosurg. Psychiatry. **54**(1), 74–76 (1991)
17. RA Lester, BH Story, Acoustic characteristics of simulated respiratory-induced vocal tremor. Am. J. Speech Lang. Pathol. **22**(2), 205–211 (2013)
18. EM Finnegan, ES Luschei, JM Barkmeier, HT Hoffman, Synchrony of laryngeal muscle activity in persons with vocal tremor. Arch. Otolaryngology-Head Neck Surg. **129**(3), 313 (2003)
19. R Collier, Laryngeal muscle activity, subglottal air pressure, and the control of pitch in speechTechnical Report Status Report on Speech Research SR-39/40, Haskins Laboratory (1974). http://www.haskins.yale.edu/sr/SR039/SR039_09.pdf (visited May, 2013)
20. JE Atkinson, Correlation analysis of the physiological factors controlling fundamental voice frequency. J. Acoust. Soc. Am. **63**(1), 211–222 (1978)
21. T Shipp, ET Doherty, P Morrissey, Predicting vocal frequency from selected physiologic measures. J. Acoust. Soc. Am. **66**(3), 678–684 (1979)
22. DK Chhetri, J Neubauer, E Sofer, DA Berry, Influence and interactions of laryngeal adductors and cricothyroid muscles on fundamental frequency and glottal posture control. J. Acoust. Soc. Am. **135**(4), 2052–2064 (2014)
23. SY Lowell, BH Story, Simulated effects of cricothyroid and thyroarytenoid muscle activation on adult-male vocal fold vibration. J. Acoust. Soc. Am. **120**(1), 386–397 (2006)
24. RA Lester, J Barkmeier-Kraemer, BH Story, Physiologic and acoustic patterns of essential vocal tremor. J. Voice. **27**(4), 422–432 (2013)
25. J Sundberg, Acoustic and psychoacoustic aspects of vocal vibrato. Quarterly Progress and Status Report 2-3, KTH. **35**, 45–68 (1994)
26. J Hanquinet, F Grenez, J Schoentgen, in *Nonlinear Analyses and Algorithms for Speech Processing*. Synthesis of disordered voices (Heidelberg Springer, 2005), pp. 231–241
27. Y Zhang, JJ Jiang, Nonlinear dynamic mechanism of vocal tremor from voice analysis and model simulations. J. Sound Vibration. **316**(1), 248–262 (2008)
28. R Fraile, M Kob, JI Godino-Llorente, N Sáenz-Lechón, VJ Osma-Ruiz, JM Gutiérrez-Arriola, Physical simulation of laryngeal disorders using a multiple-mass vocal fold model. Biomed. Signal Process. Control. **7**(1), 65–78 (2012)
29. M Kob, Physical modeling of the singing voice. PhD thesis, Fakultät für Elektrotechnik und Informationstechnik, RWTH Aachen, Logos-Verlag (2002). http://darwin.bth.rwth-aachen.de/opus3/volltexte/2002/393/pdf/Kob_Malte.pdf
30. E Loch, S Noelle, M Kob, in *Proc. of the International Conference on Acoustics, Including the 35th German Annual Conference on Acoustics*. An approach for stable calculation of vocal fold oscillation (NAG/DAGA Rotterdam, 2009), pp. 1715–1717
31. PH Dejonckere, M Kob, Pathogenesis of vocal fold nodules: new insights from a modelling approach. Folia Phoniatrica et Logopaedica. **61**(3), 171–179 (2009)
32. D Wong, MR Ito, NB Cox, IR Titze, Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases. J. Acoust. Soc. Am. **89**(1), 383–394 (1991)
33. IR Titze, BH Story, Rules for controlling low-dimensional vocal fold models with muscle activation. J. Acoust. Soc. Am. **112**, 1064–1076 (2002)
34. KA Kendall, RJ Leonard, Interarytenoid muscle botox injection for treatment of adductor spasmodic dysphonia with vocal tremor. J. Voice. **25**(1), 114–119 (2011)
35. N Quinn, Parkinsonism–recognition and differential diagnosis. Br. Med. J.(BMJ). **310**, 447–452 (1995)
36. J Jankovic, Parkinson?s disease: clinical features and diagnosis. J. Neurol. Neurosurg. Psychiatry. **79**(4), 368–376 (2008)
37. P Mermelstein, Articulatory model for the study of speech production. J. Acoust. Soc. Am. **53**(4), 1070–1082 (1973)
38. WT Fitch, J Giedd, Morphology and development of the human vocal tract: A study using magnetic resonance imaging. J. Acoust. Soc. Am. **106**, 1511–1522 (1999)
39. BH Story, IR Titze, Parameterization of vocal tract area functions by empirical orthogonal modes. J. Phonetics. **26**(3), 223–260 (1998)
40. AG Shaikh, K Miura, LM Optican, S Ramat, RM Tripp, DS Zee, Hypothetical membrane mechanisms in essential tremor. J. Translational Med. **6**(1), 68–78 (2008)
41. C Dromey, P Warrick, J Irish, The influence of pitch and loudness changes on the acoustics of vocal tremor. J. Speech Lang. Hear. Res. **45**(5), 879–890 (2002)
42. IR Titze, Nonlinear source-filter coupling in phonation: Theory. J. Acoust. Soc. Am. **123**(5), 2733–2749 (2008)