



# The LF Model in the Frequency Domain for Glottal Airflow Modelling without Aliasing Distortion

Christer Gobl

Phonetics and Speech Laboratory, School of Linguistic, Speech and Communication Sciences  
Trinity College Dublin, Ireland

cegobl@tcd.ie

## Abstract

Many of the commonly used voice source models are based on piecewise elementary functions defined in the time domain. The discrete-time implementation of such models generally causes aliasing distortion, which make them less useful for certain applications. This paper presents a method which eliminates this distortion. The key component of the proposed method is the frequency domain description of the source model. By deploying the Laplace transform and phasor arithmetic, closed-form expressions of the source model spectrum can be derived. This facilitates the calculation of the spectrum directly from the model parameters, which in turn makes it possible to obtain the ideal discrete spectrum of the model given the sampling frequency used. This discrete spectrum is entirely free of aliasing distortion, and the inverse discrete Fourier transform is used to compute the sampled glottal flow pulse. The proposed method was applied to the widely used LF model, and the complete Laplace transform of the model is presented. Also included are closed-form expressions of the amplitude spectrum and the phase spectrum for the calculation of the LF model spectrum.

**Index Terms:** voice source, glottal airflow, frequency domain LF model, aliasing, Laplace transform, phasor

## 1. Introduction

Voice source models are generally defined in the time domain in terms of a set of piecewise elementary functions. Such models include the Liljencrants-Fant model (LF model) [1], the Fujisaki-Ljungqvist model [2], the models proposed by Rosenberg [3] and many others, e.g., [4-10].

The voice source signal is a continuous-time signal, and hence the functions that are used to model this signal are continuous in time. When implemented in a discrete-time system, this is typically done by sampling the functions, i.e. by replacing the continuous variable  $t$  with  $nT_s$ , where  $n = 0, 1, 2, \dots$  and  $T_s$  is the sampling period (the reciprocal of the sampling frequency,  $1/F_s$ ). The sampling process will limit the bandwidth of the sampled waveform to the Nyquist frequency, i.e. to half the sampling frequency. Apart from this obvious fact, there will, however, also be aliasing distortion introduced, as the source spectrum is not bandlimited. Thus, the implemented discrete-time version of these continuous-time models will have a distorted frequency spectrum, even within the bandwidth determined by the sampling frequency.

At very high sampling frequencies, this is generally a relatively minor issue, but in situations where a low sampling frequency is required, it can lead to significant distortion, particularly for voices with a high fundamental frequency.

This fold-over distortion, where higher frequencies appear as lower frequencies, will manifest as spurious frequency components in the spectrum, except in the case where the sampling frequency happens to be a multiple of the fundamental frequency. In this case, no such components are introduced since each alias component will have a frequency which is the same as one of the actual harmonics. Nevertheless, the amplitude and phase of the harmonics will be altered. This latter effect is perceptually less disturbing compared to the introduction of inharmonic components in the spectrum. As  $f_0$  is constantly changing in speech and singing, the nature and degree of the distortion is also changing. This is clearly not ideal if high precision analysis and modelling of the voice source is the objective.

What can be done to avoid this distortion? Since the amplitude of the source model spectrum decreases with increasing frequency, it can be reduced by employing oversampling of the model waveform. This oversampled waveform is then lowpass filtered (to reduce the amplitudes of components above the Nyquist frequency) and subsequently downsampled to the required sampling frequency. However, this approach may not always be the best option since a high level of oversampling may be required to achieve optimal results. Furthermore, it cannot fully cancel the aliasing and the filtering involved could potentially introduce other problems.

The standard approach to the discretisation of a continuous-time signal is to suppress frequency components above the Nyquist frequency prior to the sampling of the signal. This is done by applying an anti-aliasing filter (i.e. lowpass filter) in the time domain. Based on these general principles, a technique is presented in [11] for analytically applying anti-aliasing filtering to pulses described by piecewise polynomial curves. The filtering is applied in the time domain prior to the sampling of the curves. This framework was also adopted and developed further in [12] and [13] for the LF model and the Fujisaki-Ljungqvist model respectively. By applying a set of anti-aliasing functions, the source models are modified in the time domain prior to the sampling. This effectively removes alias frequency components, but also causes severe attenuation of the amplitudes of the original spectral components [13]. Hence, postfiltering is required to restore the correct amplitudes. However, this postfiltering may in itself introduce distortion and lead to perceptual artefacts, as suggested in [12].

The current paper proposes an alternative method, where the anti-aliasing filtering is carried out directly in the frequency domain. This frequency domain processing is very simple and does not require any special anti-aliasing functions. Furthermore, it eliminates the anti-aliasing altogether and consequently no postfiltering is necessary since the correct spectrum is maintained up to the Nyquist frequency. However,

for this to be feasible, a parametric representation of the model in the frequency domain is required.

## 2. Proposed method for aliasing-free voice source modelling

The proposed method for deriving discrete-time glottal flow waveforms from a continuous-time glottal flow model, without introducing aliasing distortion, is described below in terms of five main steps.

In the first step, the Laplace transform is obtained for each part of the piecewise function defining the model. Due to the linearity property of the Laplace transform, the complete Laplace transform is the sum of the individual transforms.

Secondly, for each of the Laplace transforms, the frequency response is derived by evaluating the transform on the imaginary axis, i.e. by substituting the complex variable  $s$  in the Laplace transform with  $j\omega$ . The frequency response obtained is equivalent to the frequency spectrum of the model. We get the amplitude response by calculating the absolute value of the complex frequency response, whereas the phase response is given by the angle determined by the real and imaginary parts of the frequency response.

In the third step, closed-form expressions for the overall model spectrum are derived. These expressions enable the calculation of the spectrum directly from the model parameters, which is crucial to the objective of the proposed method. To do this, basic phasor arithmetic is employed, and each component of the spectrum is treated as a phasor. The process is illustrated in Fig. 1, showing the result of adding two phasors, where the length of the phasor represents the amplitude ( $A$ ) and the angle (relative to the real axis) represents the phase ( $\phi$ ). The main formulas are shown in (1) and (2) for the amplitude and phase respectively.

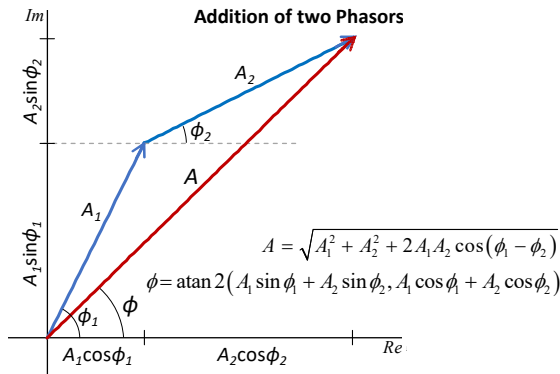


Figure. 1: The addition of two phasors, where  $A$  represents the amplitude response and  $\phi$  the phase response of two parallel (added) components.

$$A = \sqrt{(A_1 \cos \phi_1 + A_2 \cos \phi_2)^2 + (A_1 \sin \phi_1 + A_2 \sin \phi_2)^2} = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos(\phi_1 - \phi_2)} \quad (1)$$

$$\phi = \text{atan2}(A_1 \sin \phi_1 + A_2 \sin \phi_2, A_1 \cos \phi_1 + A_2 \cos \phi_2) \quad (2)$$

Note that the atan2 function is used here to ensure that the phase angles are always correct. The definition of the atan2 function is shown in (3).

$$\text{atan2}(y, x) = \arctan\left(\frac{y}{x}\right) \cdot \text{sgn}(x)^2 + \pi \left(\frac{1 - \text{sgn}(x)}{2}\right) (1 + \text{sgn}(y) - \text{sgn}(y)^2) \quad (3)$$

In the fourth step, the ideal discrete spectrum of the model is derived from the true model spectrum. This spectrum is a sampled version of the true spectrum up to the Nyquist frequency. We get the discrete frequency samples by replacing the continuous frequency variable  $f$  in the closed-form expressions of the model with  $kF_s/N$  for  $k = 0, 1, 2, \dots, N/2 - 1$ , where  $N$  is an integer which needs to be greater than the number of samples required to model the glottal pulse. As this is a discrete spectrum of a real-valued signal, the complex spectrum should be *even symmetric* and periodic with period  $F_s$ . Consequently, the frequency samples for  $k$  between  $N/2$  and  $N - 1$  can be derived from the sample values between 0 and  $N/2 - 1$ .

Once we have the frequency samples defining the ideal spectrum, the glottal pulse corresponding to this spectrum is obtained by computing the inverse discrete Fourier transform (IDFT). Note that the normalisation factor of the IDFT should be  $F_s/N$  for the correct scaling of the waveform samples.

The process described in the last three steps can also be carried out for multiple glottal pulses by applying the phasor arithmetic for the individual pulses, i.e. by treating the spectrum of each glottal pulse as individual phasors. To each subsequent pulse, a linear phase component is added, which corresponds to the duration of its delay in relation to the previous pulses. For example, if the duration of the first pulse is  $T_0$ , the additional phase component that should be added to the phase spectrum of the second pulse is  $-2\pi f T_0$ .

## 3. The parameters of the LF model

For the application of the proposed method, the LF model [1] was chosen since it is arguably the most widely used parametric glottal flow model. It has been used in many different types of voice source studies, involving analysis, synthesis and perception (e.g., [14-34]).

The LF model is a model of differentiated glottal flow, and the piecewise function defining the pulse is made up of two parts according to (4), see also the lower panel of Fig. 2.

$$U_g'(t) = \begin{cases} E_0 e^{\alpha t} \sin \omega_g t & 0 \leq t < t_e \\ -\frac{E_e}{\varepsilon T_a} (e^{-\varepsilon(t-t_e)} - e^{-\varepsilon T_b}) & t_e \leq t < t_c \end{cases} \quad (4)$$

The first part models the flow derivative during the open phase of the glottal cycle using a sine function, which increases exponentially in amplitude. The duration of the open phase is  $T_e$  and it terminates at timepoint  $t_e$  when the amplitude of the waveform is  $-E_e$ , the negative amplitude of the main glottal excitation. The second part models the flow derivative after the main excitation, during the return phase of the pulse. This is essentially an exponential function, which terminates at timepoint  $t_c$ . The duration of the return phase is  $T_b$  (see Fig. 2).

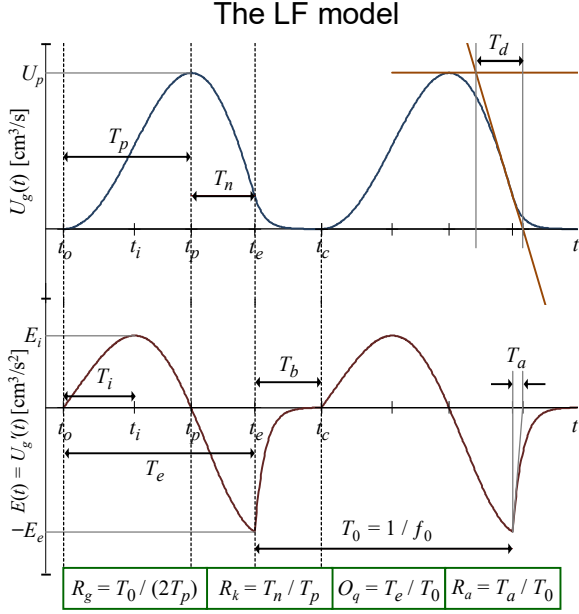


Figure 2. Two LF model pulses and parameter definitions. Flow derivative (bottom) and corresponding glottal flow (top).

The shape of the LF model pulse is often described in terms of the so-called  $R$ -parameters  $R_g$ ,  $R_k$  and  $R_a$  (see Fig. 2). The  $R$ -parameters and  $E_e$  are commonly referred to as the “LF-parameters”. However, as can be seen in (4), apart for  $E_e$ , they are not the actual LF parameters. Note further that  $E_0$  and  $T_a$  are auxiliary parameters.  $E_0$  is determined by  $E_e$ ,  $\alpha$ ,  $T_e$ , and  $\omega_g$ , and  $T_a$  is determined by  $\varepsilon$  and  $T_b$  according to (5) and (6) respectively.

$$E_0 = -E_e e^{-\alpha T_e} / \sin \omega_g T_e \quad (5)$$

$$T_a = (1 - e^{-\varepsilon T_b}) / \varepsilon \quad (6)$$

Hence, the parameters that specify the shape of the open phase are  $E_e$ ,  $T_e$ ,  $\omega_g$  and  $\alpha$ . The return phase is defined by and  $E_e$ ,  $\varepsilon$  and  $T_b$ , and the total duration of the pulse is  $T_e + T_b$ .  $T_e$  and  $\omega_g$  are easily derived from the  $R$ -parameters and  $f_0$ :  $T_e = (1 + R_k) / (2R_g f_0)$  and  $\omega_g = 2\pi R_g f_0$ . However,  $\alpha$  and  $\varepsilon$  have to be estimated iteratively:  $\varepsilon$  is derived from  $T_a$  and  $T_b$ , and  $\alpha$  is implicitly determined by the other parameters and from the LF model requirement of a given net flow gain, which is typically set to 0. Thus, the area of the positive part of the flow derivative should equal the area of the negative part.

In [17], the LF model is presented with a transformed set of parameters, which includes the global waveshape parameters  $R_d = 0.11^{-1} f_0 U_p / E_e$ . For the conversion of  $R_d$  data into actual LF parameter values, see [35].

#### 4. The LF model in the frequency domain

In developing the frequency domain representation of the LF model, it is convenient to view the LF pulse as the non-causal response to an impulse with amplitude  $-E_e$  at timepoint  $t_e$  (for a similar treatment see [36]). We therefore define  $t = 0$  as the timepoint of the main excitation ( $t_e$ ) rather than the timepoint of the glottal opening ( $t_0$ ), which is normally the case, see (4) and Fig. 2. The first part of the model (the open phase) is the

anti-causal response and the second part (the return phase) is the causal response.

By inserting (5) and (6) into (4) and defining  $t_e = 0$ , we recast (4) into the expressions in (7):

$$U_g'(t) = \begin{cases} -E_e \frac{e^{\alpha t} \sin(\omega_g(t + T_e))}{\sin \omega_g T_e} & -T_e \leq t < 0 \\ -E_e \frac{(e^{-\varepsilon t} - e^{-\varepsilon T_b})}{(1 - e^{-\varepsilon T_b})} & 0 \leq t < T_b \end{cases} \quad (7)$$

Thus, the LF model open phase can be viewed as a time-reversed, phase-shifted, exponentially damped sine function, which is truncated after a time interval of  $T_e$ . Note that the Laplace transform for an anti-causal response can be obtained by replacing  $s$  with  $-s$  in the transform of the corresponding causal response. The Laplace transform of the LF model open phase is shown in (8), where  $u(t)$  is the unit step function.

$$H_{LF_{\text{open}}}(s) = \mathcal{L} \left\{ \left[ \frac{-E_e e^{\alpha t} \sin(\omega_g(t + T_e))}{\sin \omega_g T_e} \right] [u(-t) - u(-t - T_e)] \right\} = E_e \frac{s - \alpha + \omega_g \csc \omega_g T_e (\cos \omega_g T_e - e^{(s-\alpha)T_e})}{[s - (\alpha + j\omega_g)] \times [s - (\alpha - j\omega_g)]} \quad (8)$$

The return phase is essentially an exponential function to which a constant is added. This term is needed for the waveform to equal 0 at  $t_c$  (i.e. at the end of the pulse). The Laplace transform of the LF model return phase is shown in (9).

$$H_{LF_{\text{ret}}}(s) = \mathcal{L} \left\{ \left[ -\frac{E_e}{1 - e^{-\varepsilon T_b}} (e^{-\varepsilon t} + e^{-\varepsilon T_b}) \right] [u(t) - u(t - T_b)] \right\} = -E_e \frac{s - \varepsilon e^{-\varepsilon T_b} (1 - e^{-\varepsilon T_b})^{-1} (1 - e^{-s T_b})}{s(s + \varepsilon)} \quad (9)$$

Since the LF pulse is the sum of the open phase and the return phase, the complete transform for the LF model is

$$H_{LF}(s) = H_{LF_{\text{open}}}(s) + H_{LF_{\text{ret}}}(s). \quad (10)$$

Next,  $s$  is replaced by  $j\omega$  in (10) to derive the corresponding frequency response, which is equivalent to the frequency spectrum of the model. The absolute value of  $H_{LF}(j\omega)$  is the amplitude spectrum:  $A_o(f)$  and  $A_r(f)$  for the open phase (11) and the return phase (15) respectively. We get the phase spectrum from the angle determined by the real and imaginary parts of the transform:  $\phi_o(f)$  and  $\phi_r(f)$  for the open phase (12) and return phase (16) respectively.

$$A_o(f) = |H_{LF_{\text{open}}}(j\omega)| = E_e \times \frac{\sqrt{[\text{Re}_1(f)]^2 + [\text{Im}_1(f)]^2}}{\sqrt{\alpha^2 + (2\pi f - \omega_g)^2} \sqrt{\alpha^2 + (2\pi f + \omega_g)^2}} \quad (11)$$

$$\phi_o(f) = \angle H_{LF_{\text{open}}} = \text{atan} 2(\text{Im}_1(f), \text{Re}_1(f)) - \text{atan} 2(2\pi f - \omega_g, -\alpha) - \text{atan} 2(2\pi f + \omega_g, -\alpha) \quad (12)$$

where  $\text{Re}_1$  and  $\text{Im}_1$  are given by (13) and (14).

$$\operatorname{Re}_1(f) = \omega_g \left( \cot \omega_g T_e - e^{-\alpha T_e} \frac{\cos 2\pi f T_e}{\sin \omega_g T_e} \right) - \alpha \quad (13)$$

$$\operatorname{Im}_1(f) = 2\pi f - \omega_g e^{-\alpha T_e} \frac{\sin 2\pi f T_e}{\sin \omega_g T_e} \quad (14)$$

$$A_r(f) = |H_{LFret}| = E_e \times \frac{\sqrt{[\operatorname{Re}_2(f)]^2 + [\operatorname{Im}_2(f)]^2}}{2\pi f \sqrt{\varepsilon^2 + (2\pi f)^2}} \quad (15)$$

$$\phi_r(f) = \angle H_{LFret} = \operatorname{atan}2(\operatorname{Im}_2(f), \operatorname{Re}_2(f)) - \operatorname{atan}2(2\pi f, \varepsilon) - \frac{\pi}{2} \quad (16)$$

where

$$\operatorname{Re}_2(f) = \frac{\varepsilon e^{-\varepsilon T_b}}{1 - e^{-\varepsilon T_b}} (1 - \cos 2\pi f T_b) \quad (17)$$

$$\operatorname{Im}_2(f) = \frac{\varepsilon e^{-\varepsilon T_b}}{1 - e^{-\varepsilon T_b}} \sin 2\pi f T_b - 2\pi f \quad (18)$$

By applying the phasor arithmetic described in Section 2, we get the complete LF model spectrum shown in (19) and (20), where  $\operatorname{atan}2$  is defined according to (3):

$$A_{LF}(f) = \sqrt{A_o^2 + A_r^2 + 2A_o A_r \cos(\phi_o - \phi_r)} \quad (19)$$

$$\phi_{LF}(f) = \operatorname{atan}2(A_o \sin \phi_o + A_r \sin \phi_r, A_o \cos \phi_o + A_r \cos \phi_r) \quad (20)$$

Hence, the LF model spectrum can be derived from the model parameters by using the closed-form expression in (11) to (20). To shift the pulse to start at timepoint  $t_o = 0$  rather than at  $t_e$ , a phase delay of  $-2\pi f T_e$  should be added to (20).

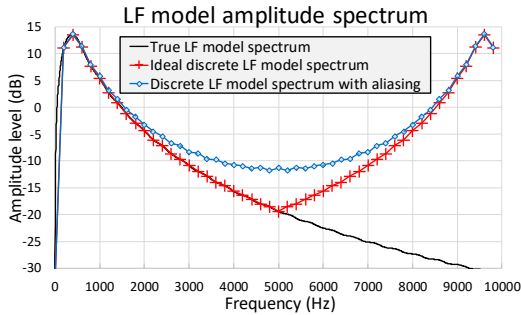


Fig. 3: Amplitude spectrum of the LF model showing the true spectrum (black), ideal discrete spectrum (red) and aliased discrete spectrum (blue).  $F_s = 10$  kHz.

As described in Section 2, we can get the ideal discrete spectrum by sampling the LF model spectrum up to the Nyquist frequency. Fig. 3 illustrates this process for the amplitude spectrum using a sampling frequency of 10 kHz. The red crosses show the frequency samples obtained from the true LF model spectrum (black line). The corresponding spectrum of the sampled LF pulse is shown in blue. Compared to the ideal spectrum, the elevated levels for frequencies approaching the Nyquist frequency reflect the aliasing distortion introduced.

The return phase of the LF model is meant to produce an increase in the slope of model spectrum: a spectral slope of about  $-12$  dB per octave is expected above the frequency approximately given by  $F_a = 1/(2\pi T_a)$ . In the example in Fig. 3,  $F_a = 1,200$  Hz. It is clear that the spectral slope is

not as steep as expected in the discrete spectrum due to the aliasing distortion.

By computing the IDFT of the ideal discrete spectrum, we get the corresponding sampled LF pulse. Fig. 4 compares two LF pulses: the pulse in red was derived using the proposed method and the pulse in blue is the corresponding sampled LF waveform. The difference between the two (green line) shows that the largest differences are found close to the main excitation and oscillate between positive and negative values. Further away from the excitation, the differences are very small, and time domain measures such as the  $R$ -parameters are only marginally affected. The extent to which they may be different for a wide range of LF pulses would, however, require further investigation.

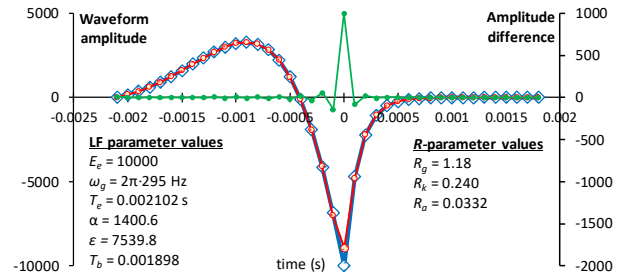


Fig. 4: LF pulse using the proposed method (red) compared to the corresponding sampled LF pulse (blue). The difference between the two is shown by the green line.

If several pulses are concatenated in the way outlined in Section 2, we get a harmonic spectrum, which is free from spurious frequency components and where the amplitude and phase of the harmonics are undistorted.

## 5. Conclusions

A method is presented which permits the implementation of time domain glottal flow models free of aliasing distortion. A key component of the method is the development of a frequency domain representation of the time domain model. Such a representation is presented here for the LF glottal flow model. Closed-form expressions for the amplitude and phase spectrum are derived, which facilitate direct calculation of the spectrum from the model parameters.

Compared to other methods, this method eliminates aliasing distortion completely, it does not require special anti-aliasing functions, and postfiltering to correct spectral amplitude levels is not necessary. Furthermore, it is conceptually intuitive and, once the frequency domain representation has been obtained, it is relatively straightforward to implement.

The method has the potential to improve the usability of glottal flow models, particularly in applications where the sampling frequency needs to be relatively low, and we are currently planning to incorporate this version of the LF model in our analysis and synthesis system GlórCáil [37].

## 6. Acknowledgements

This research was carried out in the ABAIR-Róibóglór projects, supported by the Irish Govt. Department of Tourism, Culture, Arts, Gaeltacht, Sport and Media, with funding from the National Lottery, supporting An Stráitéis 20 bliain don Ghaeilge 2010-2030.

## 7. References

- [1] G. Fant, J. Liljencrants, and Q. Lin, "A four-parameter model of glottal flow," *STL-QPSR*, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, vol. 26 (4), pp. 1-13, 1985.
- [2] H. Fujisaki and M. Ljungqvist, "Proposal and evaluation of models for the glottal source waveform," *Proceedings IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Tokyo, Japan, pp. 1605-1608, 1986.
- [3] A. E. Rosenberg, "Effect of glottal pulse shape on the quality of natural vowels," *Journal of the Acoustical Society of America*, vol. 49, no. 2B, pp. 583-590, 1971.
- [4] T. V. Ananthapadmanabha, "Acoustic analysis of voice source dynamics," *STL-QPSR*, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, vol. 25 (2-3), pp. 1-24, 1984.
- [5] G. Fant, "Glottal source and excitation analysis," *STL-QPSR*, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, vol. 20 (1), pp. 85-107, 1979.
- [6] G. Fant, "Vocal source analysis – a progress report," *STL-QPSR*, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, vol. 20 (3-4), pp. 31-53, 1979.
- [7] P. Hedelin, "A glottal LPC-vocoder," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, San Diego, California, pp. 21-24, 1984.
- [8] Y.-L. Shue and A. Alwan, "A new voice source model based on high-speed imaging and its application to voice source estimation," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, Texas, pp. 5134-5137, 2010.
- [9] H. W. Strube, "Determination of the instant of glottal closure from the speech wave", *Journal of the Acoustical Society of America*, vol. 56, no. 5, pp. 1625-1629, 1974.
- [10] R. Veldhuis, "A computationally efficient alternative for the Liljencrants-Fant model and its perceptual evaluation," *Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 566-571, 1998.
- [11] P. H. Milenkovic, "Voice source model for continuous control of pitch period," *Journal of the Acoustical Society of America*, vol. 93, no. 2, pp. 1087-1096, 1993.
- [12] H. Kawahara, K.-I. Sakakibara, H. Banno, M. Morise, T. Toda, and T. Irino, "Aliasing-free implementation of discrete-time glottal source models and their applications to speech synthesis and F0 extractor evaluation," *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Hong Kong, pp. 520-529, 2015.
- [13] K. Hideki Kawahara, K.-I. Sakakibara, M. Morise, H. Banno, T. Toda, and T. Irino, "A new cosine series antialiasing function and its application to aliasing-free glottal source models for speech and singing synthesis", *INTERSPEECH 2017*, Stockholm, Sweden, pp. 1358-1362, 2017.
- [14] J. P. Cabral, S. Renals, J., Yamagishi, and K. Richmond, "HMM-based speech synthesiser using the LF-model of the glottal source," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4704-4707, 2011.
- [15] R. Carlson, G. Fant, C. Gobl, B. Granström, I. Karlsson, and Q. Lin, "Voice source rules for text-to-speech synthesis," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Glasgow, UK, vol. 1, pp. 223-226, 1989.
- [16] A. del Pozo and S. Young, "The linear transformation of LF glottal waveforms for voice conversion," *INTERSPEECH 2008*, Brisbane, Australia, pp. 1457-1460, 2008.
- [17] G. Fant, "The LF-model revisited. Transformations and frequency domain analysis," *STL-QPSR*, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, vol. 36 (2-3), pp. 119-156, 1995.
- [18] G. Fant, "The voice source in connected speech," *Speech Communication*, vol. 22, pp. 125-139, 1997.
- [19] G. Fant, A. Kruckenberg, J. Liljencrants, and M. Båvegård, "Voice source parameters in continuous speech. Transformation of LF-parameters," *Proceedings of the International Conference on Spoken Language Processing*, Yokohama, pp. 1451-1454, 1994.
- [20] C. Gobl, "The Voice Source in Speech Communication: Production and Perception Experiments Involving Inverse Filtering and Synthesis," PhD thesis, KTH, Stockholm, Sweden, 2003.
- [21] C. Gobl and A. Ni Chasaide, "Acoustic characteristics of voice quality," *Speech Communication*, vol. 11, nos. 4-5, pp. 481-490, 1992.
- [22] C. Gobl and A. Ni Chasaide, "The role of voice quality in communicating emotion, mood and attitude", *Speech Communication*, vol. 40, pp.189-212, 2003.
- [23] C. Gobl and J. Mahshie, "Inverse filtering of nasalized vowels using synthesized speech," *Journal of Voice*, vol. 27, no. 2, pp. 155-169, 2013.
- [24] I. Karlsson, "Glottal waveform parameters for different speaker types," *Proceedings of 7 FASE Symposium*, Edinburgh, pp. 225-231, 1988.
- [25] I. Karlsson, "Voice source dynamics for female speakers," *Proceedings of the International Conference on Spoken Language Processing*, Kobe, Japan, pp. 225-231, 1990.
- [26] A. Ni Chasaide and C. Gobl, "Contextual variation of the vowel voice source as a function of adjacent consonants," *Language and Speech*, vol. 36, pp. 303-330, 1993.
- [27] A. Ni Chasaide, A. and C. Gobl, "Decomposing linguistic and affective components of phonatory quality," *Proceedings of the 8th International Conference on Spoken Language Processing, INTERSPEECH 2004*, Jeju Island, Korea, pp. 901-904, 2004.
- [28] J. B. Pierrehumbert, "A preliminary study of the consequences of intonation for the voice source," *STL-QPSR*, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, vol. 30 (4), pp. 23-36, 1989.
- [29] C. Ryan, A. Ni Chasaide, and C. Gobl, "Voice quality variation and the perception of affect: continuous or categorical?" *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS-15)*, Barcelona, Spain, pp. 2409-2412, 2003.
- [30] R. van Dinter, R. Veldhuis, and A. Kohlrausch, "Perceptual aspects of glottal-pulse parameter variations," *Speech Communication*, vol. 46, no. 1, pp. 95-112, 2005.
- [31] I. Yanushevskaya, C. Gobl, and A. Ni Chasaide, "Voice quality in affect cueing: does loudness matter?" *Frontiers in Psychology*, vol. 4, article 335, pp. 1-14, 2013.
- [32] I. Yanushevskaya, C. Gobl, and A. Ni Chasaide, "Cross-language differences in how voice quality and f0 contours map to affect," *Journal of the Acoustical Society of America*, vol. 144, no. 5, pp. 2730-2750, 2018.
- [33] S. Huber and A. Roebel, "On the use of voice descriptors for glottal source shape parameter estimation," *Computer Speech & Language*, vol. 28, no. 5, pp. 1170-1194, 2014.
- [34] S. Huber, A. Roebel, and G. Degottex, "Glottal source shape parameter estimation using phase minimization variants," *INTERSPEECH 2012*, Portland, Oregon, pp. 1644-1647, 2012.
- [35] C. Gobl, "Reshaping the transformed LF model: generating the glottal source from the waveshape parameter  $R_d$ ," *INTERSPEECH 2017*, Stockholm, Sweden, pp. 3008-3012, 2017.
- [36] B. Doval, C. d'Allesandro, and N. Henrich, "The voice source as a causal/anticausal linear filter," *Proceedings of the ISCA VOQUAL'03 Workshop on Voice Quality: Functions, Analysis and Synthesis*, Geneva, Switzerland, pp. 15-20, 2003.
- [37] A. Murphy, I. Yanushevskaya, A. Ni Chasaide, and C. Gobl, "Testing the GlórCáil System in a Speaker and Affect Voice Transformation Task," *Proc. 10th International Conference on Speech Prosody*, Tokyo, Japan, pp. 950-954, 2020.