A STUDY OF THE TWO-MASS MODEL IN TERMS OF ACOUSTIC PARAMETERS

Denisse Sciamarella and Christophe d'Alessandro

LIMSI-CNRS, BP 133, F91403, Orsay France

ABSTRACT

We present a study of the acoustic source parameters describing glottal flow waveforms generated by the two-mass model for vocal fold oscillations. Numerical measurements of the acoustic parameters as a function of model parameters are presented. Conclusions are drawn from these results, concerning the correlations between acoustic parameters, the effect of the vocal tract, the different regimes of oscillation of the model, and the possibility of reproducing laryngeal mechanisms with the model.

1. INTRODUCTION

The glottal source has been described using both, signal and production models. Signal models basically characterize one cycle of the derivative of the glottal flow by using as few as five acoustic parameters (e.g. Liljencrants and Fant model [1]). The qualitative description of pressed, normal, breathy, relaxed or flow voice is possible in terms of these parameters. Production models, instead, describe the vocal fold system in terms of a mechanical oscillator. In the first and most widely known production model (developed by Ishizaka and Flanagan [2]) each vocal cord is approximated by a self-oscillating system composed of two stiffness-coupled masses. This model has been recently revisited by [3]. Such models can take into account subtle features that are not reproduced by a signal model. They also allow to study the mechanisms responsible for the behavior of the source, which is extremely difficult to analyse experimentally. The price paid for this more detailed description is an increase in the number of parameters involved (more than nineteen parameters are present in [2]).

Unveiling the relationship between acoustic and physical parameters is a challenging issue in order to bring together the phenomena of voice perception and voice generation. The understanding of the variation of acoustic parameters with those of a production model would allow a deeper comprehension of the mechanisms responsible for different voice qualities. Such knowledge would eventually allow the construction of a *mixed model* that accounts for voice production in terms of a few acoustic parameters. Techniques as those presented in [4] could be used for its validation. The task is not easy because of the intricate correlation of parameters in both physical and signal models. However, there are several questions that we can attempt to answer.

For this purpose, we develop an algorithm to compute acoustic parameters from synthetic glottal flow waveforms generated with different sets of physical parameters of the two-mass model. Section 2 contains a brief description of the signal and production models we have considered, and describes the way in which acoustic parameters are estimated. Section 3 considers the role of the vocal tract in the calculation of acoustic parameters. Section 4 attempts a rough classification of physical parameters according to their influence on the acoustic properties of glottal flow waveforms and illustrates the possibilities of the model to reproduce the transition between laryngeal mechanisms. We present our conclusions in Section 5.

2. CALCULATION OF ACOUSTIC PARAMETERS

Phenomenological methods of voice analysis are based on processing the acoustic signal. The glottal volume velocity waveform is obtained from the speech signal by cancelling the effects of vocal tract resonances. Time-domain glottal-flow signal models work under the assumption that the glottal flow pulse has a *regular* shape, i.e. a shape qualitatively close to that of figure 1.



Fig. 1. Definition of parameters describing the glottal flow pulse (above) and its derivative (below). The fundamental period, T_0 , is a global parameter, which controls the speech melody; T_e is the duration of the open phase; T_p is the duration of the opening phase; T_a the effective duration of the return phase.

The set of acoustic parameters that we have chosen for the phenomenological description of the glottal flow waveform are: - the open quotient O_q (ranging between 0 and 1) defined as the ratio between the duration of the glottis open phase over the fundamental period $O_q = T_e/T_0$. In practice, it takes values ranging between 0.3 and 0.98. It affects mostly the low frequencies of the speech spectrum. The main perceptual correlate of open quotient is often described as voice pressure. A pressed voice corresponds to a small open quotient and a relaxed voice to a large one. - the speed quotient $S_q = T_p/(T_e - T_p)$ (ranging from 1 -symmetric glottal pulse- to 9 -highly asymmetric glottal pulse-). It conveys the degree of asymmetry of the glottal pulse. It is sometimes replaced by the asymetry coefficient ($\alpha_m = T_p/T_e$). In practice, it takes values ranging from 2 to 3.

- the amplitude of voicing A_v defined as the distance between the

minimum and maximum value of the glottal volume velocity, or alternatively, *the speed of closure* E which corresponds to the glottal volume velocity at the moment of closure. These are global parameters controlling the general amplitude of the source waveform, and thus vocal intensity.

- the return phase T_a , defined in figure 1, is a local parameter related to the abruptness of glottal closure.

On the other hand, production models for the description of voiced-sound generation compute the detailed behavior of the human vocal cords. In this article, we will consider the two-mass model proposed by Ishizaka and Flanagan in 1972 [2] in which the vocal cords are approximated by a self-oscillating source composed of two stiffness-coupes masses as shown in figure 2. The physical parameters governing the behavior of the source are the effective length of the cords (l_g) , the lower (m_1) and upper masses (m_2) , the respective thickness of the masses $(d_1 \text{ and } d_2)$, the linear and cubic spring constants $(k_1, k_2, \eta_{k1}, \eta_{k2})$, the spring constants at collision $(h_1, h_2, \eta_{h1}, \eta_{h2})$, the viscous resistances $(r_1 \text{ and } r_2)$, the cross-sectional areas at rest (A_{g01}, A_{g02}) and the subglottal pressure (P_s) . Additional parameters describe the shape of the vocal tract, which is represented as a bilateral transmission line, terminated in a radiation load equal to that for a circular piston in an infinite baffle.



Fig. 2. Schematic diagram of the two-mass approximation of the vocal cords.

The acoustic interaction between the vocal tract configuration and the glottal volume flow is strong and may entail the occurrence of oscillatory ripples in the glottal flow waveform. For this reason, a rigourous check for regularity of the glottal flow pulse is necessary before computation of the acoustic properties displayed by the two-mass model for a given set of physical parameters.

The acoustic parameters for a given set of physical parameters are computed as follows:

1. Running the model: The two-mass model is run for a given set of physical parameters. Voicing time is adjusted to a value which greatly exceeds the build-up time required for the oscillation to settle to a steady state. Glottal volume velocity U_g and its derivative U'_q are stored as a function of the iteration number.

2. Isolation of a sample of the glottal flow period: The glottal volume velocity is inspected backwards in time to search for the last greatest maximum within an interval established by the frequency range in spoken and sung voice. The iteration number j_f corresponding to this event is stored as the final instant of the sample. The iteration number corresponding to the initial instant of the sample j_i is found by inspecting the signal backwards from j_f . The next maximum that best approaches the value of $U_g[j_f]$ is stored as j_i . Next, the interval $[j_{min1}, j_{min2}] \subset [j_1, j_2]$ for which the signal is at its minimum value is computed. The interval $[j_i, j_f]$ is reset to start at $j_{min} = (j_{min1} + j_{min2})/2$. If $U_g[j_{min}] \neq 0$

(incomplete closure of the glottis) the parameter values for which this happens are recorded in a separate file.

3. Checking for a regular glottal flow waveform: We check for the existence of only one local maximum within the sample of U_g . We check if this property is fulfilled by the periods preceding the sample of U_g (the oscillations build-up phase is excluded from this verification). Similarly, we check for the existence of one local maximum and one local minimum within the sample of U'_g . This test proves to be strong enough to detect irregularities for the glottal pulse in the case of signals generated with the two-mass model. Sets of parameters yielding irregular signals are stored in a file.

4. Calculation of acoustic parameters: We inspect U'_g within $[j_i < j < j_f]$. We compute T_p by substracting the iteration number (j_1) corresponding to the first non zero value of U'_g and the iteration number (j_2) associated to the maximum of U_g . A_v is directly $U_g[j_2]$. We compute T_e from $(j_3 - j_1)$ where (j_3) corresponds to the minimum value of U'_g . E is directly $U'_g[j_3]$. Finally, T_a is computed by substracting the iteration number j for which $U'_g[j] > E/4$ and j_3 . The acoustic parameters are calculated in terms of these values following the definitions of section 2.

The algorithm of numerical measurement of acoustic parameters has been written following the above mentioned steps. The program carries out two kinds of operations:

- Running the two-mass model several times for different sets of physical parameters in which one or two parameters are iteratively varied, and computing acoustic parameters for each run. The input is the set of values of the physical parameters that are kept constant and the ranges to be covered by the parameters allowed to vary. The output is made up of a file containing the acoustic parameters corresponding to the varied physical parameters, a file containing the values of the physical parameters for which irregularities are detected, and a file containing the values of the physical parameters for which the glottis does not close.

- Producing the smooth temporal variation of one parameter of the set of physical parameters in one single run of the two mass-model. The output is a sound file for perceptive analysis of the variation of the chosen physical parameter.

3. THE EFFECT OF THE VOCAL TRACT

The measurement of the acoustic parameters from an experimental signal is carried out by cancelling the effects of vocal tract resonances. However, the degree of success of such an operation is not easy to evaluate. For this reason, it is interesting to consider in which cases the effect of the vocal is important. In this section, we consider the role of the vocal tract on the shape of the glottal pulse in terms of the two-mass model. For this purpose, we allow the algorithm to work with and without vocal tract.

In the case of 0 vocal tract segments, the network model for the synthesis of voiced sounds proposed in [2] is reduced to the circuit for the glottis directly followed by the radiation load. Following the notation in [2], the differential equations which relate the volume velocities of the system in the case of zero vocal tract segments are:

$$(R_{k1} + R_{k2})|U_g|U_g + (R_{v1} + R_{v2})U_g + (L_{g1} + L_{g2})\frac{dU_g}{dt} + L_R(\frac{dU_g}{dt} - \frac{dU_R}{dt}) - P_s = 0$$
$$L_R\frac{d}{dt}(U_R - U_g) + R_RU_R = 0$$

where U_g is the glottal volume velocity, t is time and the coefficients are defined exactly as in [2].

The absence of vocal tract is a guarantee for regularity of the glottal flow pulse within a large range of variation of the physical parameters. A departure from regular behavior only becomes likely if the model parameters are given values close to the bounds of the region of allowed oscillations, or close to a bifurcation of the dynamical system. As we will see in section 4, the calculation of acoustic parameters in the absence of vocal tract is a good scenario to detect transitions between distinct regimes of the dynamical system of voice production, .

Exemples of the effect of the vocal tract on the glottal flow waveform are given in figure 3, which contains the synthetic glottal pulse for different configurations of the vocal tract. The values of the physical parameters are set at the typical glottal condition.



Fig. 3. : (a) Glottal volume velocity in cm^3/s without vocal tract (full line), with a 4-segment uniform vocal tract (dashed line), with vocal tract as in vowel /u/ (points). (b) Glottal flow derivative in m^3/s^2 corresponding to (a). (c) Glottal volume velocity in cm^3/s with vocal tract as in vowel /a/ (full line), with vocal tract as in vowel /a/ (dashed line). (d) Glottal flow derivative in m^3/s^2 corresponding to (c).

Waveforms of glottal area and fundamental frequency are almost independent of the vocal-tract shape. Figures 3(a) and (b) group the cases for which glottal flow is sufficiently regular for acoustic parameters to be calculated. Nevertheless, notice that glottal flow is strictly regular only in the absence of vocal tract. In order to calculate the acoustic parameters for the cases with vocal tract in (a) we have relaxed the condition of one single local maximum and minimum for the glottal flow derivative. However, the values of acoustic parameters do not defer greatly: Results are presented in this order: without vocal tract, with uniform vocal tract, and with vocal tract as in vowel /u/ respectively. Oq =0.59, 0.58, 0.60; Sq = 1.95, 2.32, 2.61; E = 2.07, 2.28, 2.40 m^3/s^2 ; and $Av = 614,600,560cm^3/s$. As we can see from figures 3 (c) and (d), the effect of a vocal-tract configuration as in vowels |a| and |i| becomes so strong that the calculation of acoustic parameters is no longer possible.

4. BEHAVIOR OF ACOUSTIC PARAMETERS IN TERMS OF PHYSICAL PARAMETERS

The domain of variation of the acoustic parameters as physical parameters are varied is wide enough to reproduce all kinds of glottal flow behavior. For instance, $F_0 \in [82, 1300]Hz$, $O_q \in [0.3, 0.98]$, $S_q \in [0.5, 06.0]$, $T_a \in [0, 1.2]$ msec, $A_v \in [50, 6000]$ cm^3/s , $E \in [0.1, 25]m^3/s^2$ as subglottal pressure, glottal rest area, spring constants, cord tension and viscous resistances are varied.

Experimental studies of electroglottographic signals show, under certain conditions, the existence of correlations among acoustic parameters. For instance, an inverse variation of O_q with F_0 is observed in the second laryngeal mechanism. This correlation is preserved by almost all the parameters of the model except for k_2 , k_c , d_2 and the cord tension parameter Q (which appears as a mutiplicative factor in all spring constants and divides masses and thicknesses). Another frequent correlation is the inverse variation of O_q with intensity observed in the first laryngeal mechanism, or the proportional variation of F_0 with intensity. Even though intensity depends on the ensemble of acoustic parameters, a strong correlation is observed between intensity and the parameters A_v and E. The set of physical parameters that preserve the relation $1/O_q \propto A_v \propto F_0$ is made up of $\{m_1, m_2, P_s, l_g, d_1, k_2 > k_2^{lim}, k_c \subset I\}$. The set $\{k_1, A_{g01/2}, \eta_{k1/k2}, k_2 < k_2^{lim}, k_c \subset \overline{I}\}$ produces the opposite effect.



Fig. 4. Spectrogram, variation of intensity, and variation of fundamental frequency and open quotient for a glissando sung by a tenor.

It is worth noting that the laryngeal mechanism strongly affects the way in which the acoustic parameters are correlated. The transition between mechanisms has been attributed to a sudden modification of the activity of the muscles due to a physiological limit. More recently, it has been suggested that such transitions might crop up spontaneously, as a result of a sudden qualitative change in the vibration of the vocal folds during a smooth variation of physical parameters [5]. Figure 4 shows a glissando sung by a tenor obtained by electroglottography [6]. The transition between mechanisms 1 and 2 is marked by a jump of the open quotient. It is known that such transitions occur when the subglottal pressure and the cords tension are sufficiently high. Figure 5 shows the behavior of the acoustic parameters at high Q and P_s while the coupling spring parameter k_c is varied. The evolution of the open quotient ressembles the transition from mechanism 2 to mechanisme 1 depicted in figure 4. Furthermore, the other acoustic parameters evolve according to the above mentioned experimental observations. It is very likely that the coupling spring parameter plays an important role in the transition between laryngeal mechanisms, whether they are the result of a sudden or of a slow variation of the physical parameters. In any case, we can assert that variations of the acoustic parameters, similar to those observed for the transition between laryngeal mechanisms, can be reproduced with the two-mass model.



Fig. 5. : Variation of acoustic parameters at high Q = 3 and $P_s = 40 cm H_2 O$ while the coupling spring constant k_c is varied.

Within the explored ranges, this study shows that the physical parameters whose smooth variation may induce a sudden qualitative change of the signal (and consequently in the acoustic properties of the glottal flow) are $\{k_c, k_1, k_2, A_{g01/2}, r_1, r_2, P_s\}$. These sudden changes may be regarded as manifestations of bifurcations of the underlying dynamical system, as Herzel already observed for k_c and k_2 with a simplified two-mass model [5]. In the case of the two mass model at the typical glottal condition and without vocal tract, the gross features of the bifurcations in the (k_2, k_c) plane clearly appear in figure 6 (a). The regions in gray correspond to regular glottal flow waveforms whose acoustic parameters vary smoothly. The darker region corresponds to the occurrence of incomplete glottal closure. The empty band that traverses the phonation region corresponds to an appearence of a broad hump on the rising slope of the glottal flow wave that prevents the calculation of acoustic parameters and that entails a sudden qualitative change in the values of the acoustic parameters. The variation of O_q in the (k_2, k_c) plane is presented in figure 6 (b).

5. CONCLUSIONS

In this work, we have examined some of the properties of the vocal-fold 2-mass model in terms of the acoustic parameters of glottal-flow signal models. An algorithm that measures the acoustic parameters (of glottal flow waveforms generated by the model for different sets of physical parameters) has been elaborated and employed to explore the general features of the dynamical system. In this article, we have determined the conditions under which the phenomenological description provided by the signal model can be applied to 2-mass-model generated signals for different vocal



Fig. 6. : (a) Each mark indicates a regular glottal pulse generated by a run of the two-mass model for different (k_2, k_c) values in units of $[m_1/k_cm_2]$. The dark region indicates incomplete glottal closure. (b) Variation of O_q in the (k_2, k_c) plane at a typical glottal condition for the rest of the physical parameters.

tract configurations. The effect of the vocal tract on the shape of the glottal pulse appears to be important: it would be interesting to compare these results with a perceptive study of the effect. We have found out that the range of values for the acoustic parameters covered by the 2-mass model is wide enough to reproduce all kinds of observed glottal flow behavior. In particular, we have classified the physical parameters of the model according to the correlations between acoustic parameters that they induce, and compared these correlations with those observed in electroglottographic signals. We believe that a systematic measurement of acoustic parameters provides an ideal scenario to detect bifurcations of the underlying dynamical system, since it allows the identification of the physical parameters that cause sudden qualitative changes in the glottal pulse (and of the values of the parameters for which this happens). We have shown that under certain conditions, such changes may be associated to a transition between laryngeal mechanisms.

6. REFERENCES

- Fant G. Liljencrants J. and Lin Q. A four parameter model of the glottal flow. STL-QPSR, 85 (2):1-13, 1985
- [2] K. Ishizaka and J. L. Flanagan, Synthesis of Voiced Sounds form a two-mass model of the vocal cords. Bell. Syst. Tech. J., 51, pp. 1233-1268, 1972.
- [3] X. Pelorson et al. Theoretical and experimental study of quasi steady-flow separation within the glottis during phonation. J. Acoust. Soc. Am. 96 (6), 1994.
- [4] D. Sciamarella, G. B. Mindlin. Topological structure of flows from human speech data, Phys. Rev. Letters, 82, 1450 (1999).
- [5] Herzel, Bifurcations and chaos in voice signals. App. Mech. Rev. 46 (7) 1993. J. Voice 7 (1) 1993.
- [6] N. Henrich, Thèse de Doctorat de l'Université Paris 6. Etude de la source glottique en voix parlée et chantée, 2001.