A RANDOM-WALK MODEL OF JITTER

Jean Schoentgen
Laboratory of Experimental Phonetics, Université Libre de Bruxelles, Brussels, Belgium
National Fund for Scientific Research, Belgium

ABSTRACT
The presentation concerns statistical models of jitter, the small stochastic perturbations of the glottal cycle lengths. Conventional models randomly update the cycle lengths in synchrony with glottal events. We propose instead to simulate jitter by means of stochastic difference equations. The equations we propose simulate several known properties of jitter via a set of minimal assumptions. The results agree with existing models that relate jitter to the neural input to the laryngeal muscles. The results also suggest that jitter depends on the parameters that control fundamental frequency and the synchrony of vocal fold movement.

1. INTRODUCTION

Jitter may be phenomenally defined as the small random cycle-to-cycle perturbations of the glottal cycle lengths. Jitter appears to be a genuine stochastic phenomenon, distinct from intonation, quaver or tremolo, asymmetry, period doubling or chaos. Asymmetry here is a cover term that refers to spatially asymmetric vocal fold movement. Period doubling, which may take place repeatedly, is a behaviour displayed by non-linear oscillators. Period doubling is the switch, at critical values of the parameters that control an oscillator's movement, from one periodic motion to another one whose cycle length is twice as long. Chaos, finally, designates the movement of an oscillator that is unpredictable in the long term, while the oscillator is free of any stochastic components. A typical path to chaos is repeated period doubling.

Observed properties of jitter are the following.
1) Jitter is typically comprised of between 0.1 and 1 percent of the average glottal cycle length [1,11].
2) The cycle perturbations are distributed normally [2].
3) The average perturbation is proportional to the cycle lengths, i. e. longer cycle lengths appear to be, on average, more perturbed than shorter ones [2,11].
4) Perturbations of neighbouring cycles are positively correlated, i.e. cycles longer or shorter than the average tend to occur together [3,4,5].
5) Jitter in hoarse voices is expected to differ from jitter in clear voices. Generally speaking, (some) laryngeal pathologies are expected to increase jitter. But occasionally, decreases can also be observed [6,7].
6) No mechanism is known that would update the cycle lengths in synchrony with glottal events, which means that jitter presumably is the outcome of cycle-asynchronous disturbances.

One additional assumption, which is made implicitly, is that meaningful statistics regarding jitter can be obtained from a single recording of the glottal cycle lengths. Formally, this means that the process that produces jitter is ergodic. Regarding modelling, the ergodic assumption would suggest that a system simulating jitter must be at least stationary. In practice, the relevance of this assumption has been put in doubt by those who recommend recording several speech items roughly at the same time or at different times of the day [8]. Ensemble averaging of several cycle length series, however, has never been performed to this author's knowledge.

A physiology-based model of jitter has been proposed by Titze [9]. Our objective was instead to discover the simplest model that reproduces qualitatively the properties of jitter listed above. Assuming that amplitude and phase are independent variables, we posit stochastic difference equations that describe a randomly perturbed rotator model. We solve the stochastic difference equations algebraically and discuss their implications for laryngeal jitter.

2. A RANDOM-WALK MODEL OF JITER

Formulas (1) and (2) describe the movement of a rotator in polar coordinates (r,α). If the movement was periodic, the quantity to the right of equation (2) would be a constant. But, because the movement is stochastically perturbed, equation (2) transforms into equation (3), which is an instance of a Langevin equation.

$$(1)\quad \frac{dr}{dt} = 0$$

$$(2)\quad \frac{d\alpha}{dt} = \frac{2\pi}{T}$$

$$(3)\quad \frac{d\alpha}{dt} = \frac{2\pi}{T_0} + \begin{cases} +\delta, p = 0.5 \\ -\delta, p = 0.5 \end{cases}$$

$$(4)\quad \int_0^T \left(\frac{d\alpha}{dt}\right) dt = 2\pi$$

A Langevin equation is a heuristic differential equation comprising a random variable. The average cycle length T_0 is indeed perturbed by a small constant δ whose sign is assigned randomly. Symbol p is the probability. Equation (3) is not equivalent to the conventional definition of jitter, i. e. δ is not the cycle perturbations observers are wont to measure. But, observed cycle perturbations are a collective effect of many intra-cycle (unobserved) perturbations δ. The glottal cycle length T is defined via formula (4).

To obtain solutions of equation (3), it must be transformed first into a stochastic difference equation (5) that is discrete. (An alternative would have been to turn equation (3) into a stochastic differential equation.)

$$(5)\ \alpha_{n+1} = \alpha_n + \frac{2\pi}{T_0}(\Delta t) + \begin{cases} +\Delta, p = 0.5 \\ -\Delta, p = 0.5 \end{cases}$$

Model (5) is known in the literature as a (discrete) random walk. The choice of the updating frequency of the sign of disturbance Δ is not crucial, but this frequency must be higher than the average glottal cycle frequency $2\pi/T_0$. The reason is that observed jitter is normally distributed and proportional to the cycle length. This suggests that many perturbations must take place within one glottal cycle so that the law of large numbers can take effect. Since the presentation is based on algebraic solutions, we assume that the updating frequency is equal to the sampling frequency $1/(\Delta t)$.

The stochastic progression of random walk (3) of the rotator's phase functions such that the phase's target value of 2π is reached with a variable number of temporal steps. Since the cycle length is the number of steps times the sampling interval, cycle length T varies randomly around its average T0.

Indeed, an expression of the relative jitter, can be obtained by algebraically solving model (5) by repeated substitution, inserting 2π for the phase, and isolating length T ($\sim N(\Delta t)$); w_+ and w_- are the numbers of the positive and negative intra-cycle perturbations of the oscillator's phase.

$$(6)\ \frac{T - T_0}{T_0} = \frac{1}{2\pi}\Delta\left(w_- - w_+\right)$$

Formula (6) shows that the relative jitter increases with the average cycle length and that the lengths are distributed normally. Indeed, relative jitter is proportional to the difference between the positive and negative disturbances, which increases over time. Normality is a consequence of the binomial distribution involved with the (+) and (-) steps. The difference in expression (6) is an even number since the increases and decreases occur in pairs [12].

Model (5) is not stationary. The phase drifts perpetually and, as a consequence, relative jitter depends on the average cycle length. Another problem is the lack of terms in model (5) that would explain the oscillator's control over the intra-cycle disturbances and their inter-cycle relations.

3. AN ORNSTEIN-UHLENBECK MODEL OF JITTER

The phase's random walk model can be transformed into a stationary model by inserting a deterministic term that is proportional to the difference between the rotator's phase and the reference phase of a periodic vibrator.

$$(7)\ \alpha_{n+1} = \alpha_n + \frac{2\pi}{T_0}(\Delta t) - k(\alpha_n - \frac{2\pi}{T_0}(\Delta t)n) + \begin{cases} +\Delta, p = 0.5 \\ -\Delta, p = 0.5 \end{cases}$$

The continuous analogue of model (7) is known in the literature as an Ornstein-Uhlenbeck process, which is stationary [10]. Again , we do not focus on (unobserved) phase α, but rather on the perturbed cycle lengths T. An expression of relative jitter can be obtained by algebraically solving difference equation (7), inserting 2π for the phase and isolating T.

$$(8)\ \frac{T - T_0}{T_0} = \frac{1}{2\pi}\sum_{i=0}^{N-1} \mp\Delta(1-k)^i$$

The assumption is that coupling constant, k, is positive and < 1, otherwise the intra-cycle disturbances of the rotator's phase oscillate at half the sampling frequency or increase without bounds. When constant k is approximately equal to one, the perturbations are strongly dampened. When k is equal to zero, formula (8) turns into formula (6). Formula (8) also shows that relative jitter is formally stationary, because the contribution of the past disturbances becomes negligible owing to the weighting term that decreases with increasing time. Absolute jitter, however, still depends on average cycle length, T_0.

Simulations show that model (7) does not give rise to a significant correlation between the perturbations of adjacent cycle lengths, even though the disturbances of the oscillator's phase are locally correlated [10]. Another undesirable property is that for models (5) or (7), the numerical values of disturbances Δ must be chosen unrealistically large. Taking into account that $\Delta = \delta(\Delta t)$, formula (6) indeed shows that for a fixed level of relative jitter, intra-cycle disturbance δ has no upper bound and that its lower bound is equal to half the observed cycle perturbation (in Hz). In practice, this means that for models (5) and (7), (unobserved) intra-cycle disturbances are always much larger than the (observed) cycle perturbations. The reason is that intra-cycle disturbances are unitary pulses. These are applied instantaneously and their amplitudes must therefore be large to transfer a fixed amount of energy.

4. A RELAXATION MODEL OF THE LOCAL DISTURBANCES

Disturbances Δ we have considered so far are unitary pulses that are not spread out in time. An alternative is to assume that the local disturbances are significantly different from zero over a finite amount of time before decaying. Henceforth, we therefore assume that a single disturbance is a decaying exponential.

$$(9)\ \Delta_n = \begin{cases} +\Delta a^n, p = 0.5 \\ -\Delta a^n, p = 0.5 \end{cases}$$

Disturbances (9) overlap. Mathematically speaking, their combined activity is expressed by means of a sum that is inserted into models (5) and (7).

$$(10)\ \alpha_{n+1} = \alpha_n + \frac{2\pi}{T_0}(\Delta t) + \sum_{i=0}^{n} \pm\Delta a^i$$

$$(11)\ \alpha_{n+1} =$$

$$\alpha_n + \frac{2\pi}{T_0}(\Delta t) - k(\alpha_n - \frac{2\pi}{T_0}(\Delta t)n) + \sum_{i=0}^{n} \pm\Delta a^i$$

The perturbed cycle lengths can be obtained algebraically. The results are the following. Formula (12) is inferred from model (10) and formula (13) from model (11). The difference is that relative jitter (13) is stationary.

$$(12)\ \frac{T - T_0}{T_0} = \frac{1}{2\pi} \sum_{n=0}^{N} \sum_{i=0}^{n} \mp\Delta a^i$$

$$(13)\ \frac{T - T_0}{T_0} = \frac{1}{2\pi} \sum_{n=0}^{N} (1-k)^n \sum_{i=0}^{n} \mp\Delta a^i$$

In formula (13), the term between the summation signs indeed evokes stationary perturbations after a transient that depends on coupling constant, k. Simulations show that cycle lengths generated by means of model (13) correlate positively, when pulse decay, a, is approximately equal to 1. This means that the pulse's decay is spread out over many cycles. This assumption agrees with the physiological model of jitter proposed by Titze [9].

As it stands, model (13) is able to simulate jitter with the properties listed in section 1. The ability of model (13) to explain glottal cycle length jitter is discussed in the following section.

5. DISCUSSION

The formalism presented here is based on the assumption that the oscillator's frequency is perturbed directly. In the case of the laryngeal vibrator, however, the fundamental frequency is controlled via a set of parameters, which include longitudinal vocal fold tension, lateral compression, mass and length. The link between the phase perturbations considered in the previous sections and the perturbations of the laryngeal control parameters can be formalised as follows, subject to the conditions that the average glottal cycle length depends smoothly on its parameters, the vibrations are stationary, and the disturbances are small. Symbols Λ are small changes of the values of the control parameters and symbols K are constants.

$$(14)\ T - T_0 = +K_1\Lambda_1 + K_2\Lambda_2 + ... + K_m\Lambda_m$$

Formula (14) shows that a linear relation is expected to exist between absolute jitter and small perturbations of the control parameters of the cycle lengths. One consequence is that the stochastic models discussed in the previous sections apply equally to disturbances of the instantaneous frequency and of the control parameters. Constants K are the gains which relate observed cycle length jitter to control jitter.

Models (7) and (13) contain a term that couples the rotator's phase to the phase of a periodic vibrator that moderates the drift of the disturbed phase and the dependence of relative jitter on the cycle length. Obviously, this periodic oscillator has no direct laryngeal counterpart. But, the standard laryngeal model is based on the assumption that a vocal fold is made up of a minimum of two coupled oscillators. This suggests rewriting model (11) as follows.

(15)

$$\alpha_{n+1} = \alpha_n + \frac{2\pi}{T_0}(\Delta t) - k(\alpha_n - \beta_n) + \sum_{i=0}^{n} \pm\Delta a^i$$

$$\beta_{n+1} = \beta_n + \frac{2\pi}{T_0}(\Delta t) - k(\beta_n - \alpha_n) + \sum_{i=0}^{n} \pm\Delta a^i$$

The simulations we have performed show that the two stochastically disturbed oscillators in formula (15) depress each other's absolute jitter, subject to the condition that the disturbances are not spatially correlated. When they are, the net effect depends on whether the correlation is positive or negative and whether the movements are in or out of phase.

At present, it is not known whether models (14) and (15) are relevant to the explanation of glottal jitter. Still, if one takes seriously the claim that pathological modifications of the texture, shape or mass of the vocal folds condition jitter, then changes of the coupling strength between spatially separate parts of the laryngeal vibrator are a possible explanation. The reason is that phase-coupling is the most common process by which oscillators synchronise their movements and the disruption of the coupling is therefore likely to enhance the expression of external disturbances (e.g. control jitter).

REFFERENCES

[1] V. L. Heiberger, Horii Y. (1982), Jitter and Shimmer in sustained phonation, in Speech and Language : Advances in Basic Research and Practice, 7, Academic Press, New York, 299-332

[2] Laver J., Hiller S., Mackenzie J., Rooney E. (1986), An acoustic screening system for the detection of laryngeal pathology, J . Phonetics, 14, 517-524

[3] Schoentgen J., De Guchteneere R. (1991), An algorithm for the measurement of jitter, Speech Comm., 10, 5-6, 533-538
[4] Schoentgen J., De Guchteneere R. (1995), Time series analysis of jitter, J. Phonetics, 23, 1-2, 189-201
[5] Schoentgen J., De Guchteneere R. (1997), Predictable and random components of jitter, Speech Comm., 21, 255-272
[6] Lieberman P. (1963), Some acoustic measures of the fundamental periodicity of normal and pathologic larynges, J. of the Acoustical Society of America, 35, 344-353
[7] Higgins M. B., Saxman J. H. (1989b), Variations in vocal frequency perturbations across the menstrual cycle, J. Voice, 3, 233-243
[8] Higgins M. B., Saxman J. H. (1989a), A comparison of intrasubject variation across sessions of three vocal frequency perturbation measures, J. of the Acoustical Society of Amercia, 86, 911-916
[9] Titze I. (1989), A model for neurological sources of aperiodicity in vocal fold vibration, J. Speech and Hearing Research, 34, 460-472
[10] Gardiner C. W. (1985), Handbook of stochastic methods, Springer Verlag, Berlin
[11] Horii Y. (1979), Fundamental frequency perturbation observed in sustained phonation, J of Speech and Hearing Research, 22, 5-19
[12] Doucet P. and Sloep P. B. (1992), Mathematical modeling in the life sciences, Ellis Horwood, New York