Neuromechanical interaction in human snoring and upper airway obstruction

LIXI HUANG,1 AND JOHN E. FFOWCS WILLIAMS2
1Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hunghom, Kowloon, Hong Kong; and 2Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom

Huang, Lixi, and John E. Ffowcs Williams. Neuromechanical interaction in human snoring and upper airway obstruction. J. Appl. Physiol. 86(6): 1759–1763, 1999.—The fact that snoring and obstructive apnea only occur during sleep means that effective neuromuscular functioning of the upper airway during sleep is vital for the maintenance of unimpeded breathing. Recent clinical studies in humans have obtained evidence demonstrating that upper airway neural receptors sense the negative pressure generated by inspiration and “trigger,” with a certain delay, reflex muscle activation to sustain the airway that might otherwise collapse. These findings have enabled us to propose a model in which the mechanics is coupled to the neuromuscular physiology through the generation of reflex wall stiffening proportional to the retarded fluid pressure. Preliminary results on this model exhibit three kinds of behavior typical of unimpeded breathing, snoring, and obstructive sleep apnea, respectively. We suggest that the increased latency of the reflex muscle activation in sleep, together with the reduced strength of the reflex, have important clinical consequences.

SNORING OCCURS in at least 15% of the adult (especially male) population and is easily recognized as an unpleasant low-frequency noise that is accompanied by the vibration of the upper airways (2, 8). Wheezing is another example in which the airway vibrates during inspiration (3). The mechanics of both have been modeled with assumptions about the instability of the airway walls, with the wall tissue behaving like a piece of rubber. Mechanical modeling of this kind has contributed much to the understanding of physiological processes, but, in the present study, we show that it is unlikely to be sufficient for the understanding of snoring and upper airway obstruction during sleep. The present study argues the importance of neuromechanical coupling in the upper airway.

Inspiration is driven by neural drive to the inspiratory musculature creating a subatmospheric intrathoracic pressure; this pressure also tends to collapse the airway walls, predominantly at the oropharyngeal level. The system appears to counter this tendency by reflexly activating palatal and tongue muscles (that enlarge the upper airways), in addition to activating dilator muscles around the collapsing segments. The effectiveness of this mechanism is reduced in sleep. The reflexes have been known to exist in animals for some time, but have only been confirmed recently in humans. Horner and his colleagues (4–6) documented such reflexes in normal awake and sleeping subjects, resulting in activation of the tongue muscles (genioglossus). These authors found that local anaesthesia of the nose palate or pharynx reduced the power of the reflex, and they deduced that neural receptors in the mucosal surface of these structures served the local intraluminal pressure in the experimental paradigm used. Subsequently, Kobayashi and his colleagues (9) were able to show that activation of the genioglossus during inspiration enlarges the retroglossal space, at least in the anteroposterior diameter.

The palatoglossus muscle pulls the soft palate downward and forward to open the retropalatal space. Obstruction of the airway at the palatal level during sleep occurs commonly in obstructive sleep apnea and snoring. A reflex, similar to that which activates genioglossus, has now been described for palatoglossal activation with negative upper airway pressure (10); the power of this reflex is also decreased by upper airway surface anesthesia.

Horner et al. (5) showed that the magnitude of reflex activation of the genioglossus increases with the magnitude of the negative upper airway pressure; it is less effective during sleep. The onset of the muscle response had a latency of 40 ms in wakefulness and as much as 110 ms in sleep. These findings have now led us to propose a lumped-parameter model to describe the stability of airflow through the narrowest section of the upper airway. Solution of the eigenvalue problem suggests that, if the reflex latency is larger than a small critical level, then the reflex mechanism can always cause flutter, depending on what mechanical damping coefficients are assigned to the upper airway. We propose that this analysis contributes to the understanding of the flutter mechanism. We are able to define conditions for stable oscillation, flutter, and collapse in the upper airway; these states, we suggest, result in unobstructed breathing, snoring, and obstructive sleep apnea, respectively.

MODEL

Details of upper airway mechanics are complicated, and theoretical models of various sophistication have been tested. For example, Gavriely and Jensen (2) proposed a lumped-parameter model that aimed at explaining the process of airway collapse; Fodil et al. (1) studied the interaction of multiple discrete elements; whereas Huang (7) recently tested a continuum model. However, in all these models and in others relating to different aspects of snoring mechanisms the factor of neuromechanical coupling has been ignored. In an attempt to isolate the role of this factor, we believe that it is sufficient and illustrative to go back to the lumped-parameter model in which the vibrating part of the...
Airway, such as the oropharynx during snoring, shown in Fig. 1, is simplified as a “piston” with mass $m$, damping coefficient $R$, and spring stiffness $K$, all measured over a unit surface area in contact with the airflow. The definitions of these parameters are for algebraic convenience and should not be confused with the modeling of a system with multiple degrees of freedom, as in Refs. 1 and 7.

Air of density $\rho_0$ flows over the piston at speed $U$. The pressure on the surface is a function of the piston displacement. If the piston moves into the passage by a small amount $\eta$, which simulates the partial collapse of the pharyngeal passage, the flow accelerates, and the steady-flow pressure perturbation ($p'$) is given by the Bernoulli equation

$$p' = -\rho_0 U^2 \eta / h$$

where $h$ is the nominal width of the passage. Other aspects of fluid mechanics may also contribute to the dynamics of the piston, but we choose to ignore them to highlight the importance of neuromuscular functions. Acting on signals sent by the neural receptors and with a certain time delay ($\Delta t$), the dilator muscle opposes the collapsing tendency by increasing the wall stiffness by an amount proportional to the negative pressure at the retarded time $p'(t - \Delta t)$. The total force that acts to restore the piston to its equilibrium position, i.e., to oppose the airflow collapse, is

$$K \eta - A p'(t - \Delta t)$$

where $A$ is a dimensionless parameter representing the strength of the reflex mechanism. There are as yet no available data for $A$, but the range of $\Delta t$ is available for muscle reflex during wakefulness and sleep (see Refs. 4–6).

The dynamics of the piston system is governed by

$$m \ddot{\eta} + R \dot{\eta} + \left[ K - \frac{\rho_0 U^2}{h} \right] \eta + A \frac{\rho_0 U^2}{h} \eta(t - \Delta t) = 0$$

where the umlaut and the overdot denote time derivatives. The terms in the equation are, respectively, the inertia force, damping, structural restoring force, negative fluid pressure, and the extra restoring force arising from the neural stiffening. Considering harmonic oscillation of angular frequency ($\omega$), $\eta = \eta_0 \exp(i \omega t)$, where $i = \sqrt{-1}$, the above equation becomes

$$-\omega^2 + i \omega + K - \frac{\rho_0 U^2}{h} + A \frac{\rho_0 U^2}{h} \exp(-i \omega \Delta t) = 0$$

which can be solved to obtain the angular frequency $\omega$, called eigen frequency. This equation can be rewritten as

$$-\omega^2 + 2i \sigma \omega + K_{\text{mech}} + K_{\text{neuro}} \exp(-i \omega \Delta t) = 0 \quad (1)$$

by the introduction of the following parameters

$$K_{\text{mech}} = \frac{K h - \rho_0 U^2}{m h}, \quad \sigma = \frac{R}{2m}, \quad K_{\text{neuro}} = \frac{A \rho_0 U^2}{m h}$$

where $K_{\text{mech}}$ represents the mechanical stiffness of the fluid-piston system, $\sigma$ is the frictional resistance, and $K_{\text{neuro}}$ is the magnitude of the stiffness induced by the neuromuscular activation. Both $K_{\text{mech}}$ and $K_{\text{neuro}}$ have the physical dimension of $\omega^2$.

Before we proceed to solve Eq. 1, it is informative to discuss the range of values relevant to the realistic situations. For people who snore or suffer from obstructive sleep apnea, the upper airway routinely collapses, either partially or completely. That means that $K_{\text{mech}}$, which takes account of the Bernoulli effect, is probably approaching zero or even becoming negative. The extent to which this is so depends on the structural stiffness $K$. For people with relatively limp upper airway walls, $K$ is small, and air passage begins to collapse as soon as the inspiration reaches a critical level; that leads on to a higher local flow speed $U$ and a diminishing value of $K_{\text{mech}}$. The neuromuscular coefficient $A$ and stiffness parameter $K_{\text{neuro}}$ decrease as sleep becomes deeper and deeper. The neural stiffness term $K_{\text{neuro}}$ also depends on the level of negative pressure experienced at the throat. At present, we have no clinical data on which an estimate of $K_{\text{neuro}}$ can be based. We must, therefore, content ourselves with the qualitative results of system behavior at high or low

---

**Fig. 1.** Single-piston model for upper airway vibration. $\eta$, Locations where snoring often originates (left). See text for diagram symbol definitions (right).
ratio of $K_{\text{neuro}}/K_{\text{mech}}$. We also have very little idea of how the damping term $\sigma$ should be specified. What we do know is that palatal snoring is typically ~35 Hz, rising to some 100 Hz if part of the pharyngeal wall is involved.

The reflex latency $\Delta t$ measured by Horner et al. (5) using a sudden drop in airway pressure provides a clue to the kind of magnitude one should expect during sleep. There is as yet no experimental evidence to support our conjecture that the neuromuscular function is linearly correlated with the perturbation pressure. The latency is ~100 ms during sleep, which is 3 to 10 times the period of the snoring oscillation cycle, depending on the type of snore. However, there may well be other neuromuscular functions at work during breathing that contribute to the dynamics of the upper airway, such as the central neural control that precedes the respiratory maneuver. The collective effect of all the neural factors may possibly produce a smaller latency than that suggested by the isolated reflex mechanism, although the latter is normally perceived as a very quick reaction.

### Eigenvalues

The solution to Eq. 1 determines the characteristic angular frequency of the system from which a judgement can be made as to whether the system is stable to small disturbances. The roots of the equation are called eigenvalues (for frequency $\omega$). Denote the real ($\omega_r$) and imaginary ($\omega_i$) parts of the complex eigen frequency by

$$\omega = \omega_r + i \omega_i$$

we have the real and imaginary parts of Eq. 1

\[ K_{\text{neuro}} e^{i \omega t} \cos (\omega_r t + \omega_i t) = (\omega_r^2 + 2i\omega_r - (K_{\text{mech}} + \omega_i^2)) (2) \]

\[ K_{\text{neuro}} e^{i \omega t} \sin (\omega_r t + \omega_i t) = 2\omega_i (\omega_r - \omega_i) (3) \]

Since $\sin^2(\omega_r t) + \cos^2(\omega_r t) = 1$, the two equations can be combined to give

$$\omega_i t = \ln \Gamma / K_{\text{neuro}} \Gamma^2$$

\[ = |(\omega_r^2 + 2i\omega_i) - (K_{\text{mech}} + \omega_i^2)|^2 + 2\omega_i (\omega_r - \omega_i)| \]

where $\Gamma$ is an amplitude function. Equation 2 alone then gives

$$\omega_r t = 2\pi n \pm \cos^{-1} \left( |(\omega_r^2 + 2i\omega_i - K_{\text{mech}} - \omega_i^2)| / \Gamma \right) (4)$$

where $n$ is any integer, and the choice of signs $\pm$ is made by that of $\sin \omega_r t$ given in Eq. 3.

Before we calculate the eigenvalues by numerical means, it is instructive to solve for a special case analytically. When the latency is very small, the time delay term $\exp(-i\omega t)$ in Eq. 1 is simply $(1 - i\omega t)$. So the eigen frequencies are then

$$\omega = i\sigma \pm (K_{\text{mech}} + K_{\text{neuro}} - \sigma^2)^{1/2}$$

The effective stiffness is the sum of mechanical and neural stiffness, $K_{\text{mech}} + K_{\text{neuro}}$, as there is almost no delay; but the damping coefficient is reduced from $\sigma$ by an amount $K_{\text{neuro}} \Delta t/2$. When the total damping coefficient becomes negative, $\sigma < 0$, amplification of disturbance occurs, namely, flutter. Part of the neuromuscular force has transformed itself from a restoring force to a negative damping, which causes instability. The energy of oscillation in this case comes from the muscle instead of the flow.

When $\Delta t$ is large, say, several times the eigen oscillation cycle, Eqs. 2 and 3 have to be solved numerically. Perhaps the easier way to find the eigenvalues is to fix $\omega$, first and solve for $\omega_r$ by canceling out $\Delta t$ from Eqs. 4 and 5. Then $\Delta t$ is found as a result instead of being assumed a priori. The final result should be a functional relationship between $\omega$ and $\Delta t$, or, simply, $\omega(\Delta t)$, for which typical graphs are shown in Fig. 2. In preparing the graphs, most dimensional quantities have been normalized by using a reference frequency ($\omega_{\text{ref}}$) defined as

$$\omega_{\text{ref}} = (K_{\text{mech}} + K_{\text{neuro}})^{1/2}$$

As shown in Fig. 2, A and B, for a given value of $\Delta t$, there are many eigen frequencies $\omega$. At least one of them has negative imaginary part, $\omega_i < 0$; those correspond to instabilities. As $\Delta t$ increases, each eigen branch, $n = 0, 1, 2, ...,$, starts with a positive $\omega_r$, which becomes negative and then approaches a line close to the neutral one, $\omega_i = 0$, as $\Delta t \to \infty$. It has to be pointed out that the level of lowest $\omega_r$ reached (i.e., the most unstable mode) depends on the mechanical damping coefficient $\sigma$, for which Fig. 2 is 0.08$\omega_{\text{ref}}$. For higher values of $\sigma$, these lines of $\omega_r(\Delta t)$ can be elevated to positive values of $\omega_r$ for all $\Delta t$. Notice that for stable modes, $\omega_r > 0$, the curves of angular frequency rise above unity, i.e., $\omega_r > \omega_{\text{ref}}$, because the neuromuscular reaction is, in fact, determined by the displacement at earlier times, which has larger vibration amplitude. The result is that the apparent stiffness is higher than a no-latency system.

Figure 2C plots the growth factor, which is the ratio of amplitude of one cycle to that of the previous cycle, against the latency expressed in terms of the number of cycles delayed. It can be seen that the curves for $n = 0, 1, 2, ...,$, roughly correspond to solutions around $n$ cycles of delay, $\omega_r t \approx 2\pi n$. Finally, for very small $\Delta t$, there is a range in which no unstable eigen frequency exists. That latency is just enough to overcome the level of mechanical damping given by $\sigma$.

### Snoring and Other Respiratory Conditions

Snoring is caused by the vibration of the upper airway. In linear analysis, such vibration is seen as the manifestation of unstable eigen vibration modes. We, therefore, identify snoring by an eigen frequency with a nonzero $\omega_r$ and negative $\omega_i$. Normal breathing, on the other hand, should correspond to the stable modes, which have positive $\omega_i$. The borderline is $\omega_i = 0$, for which Eqs. 2 and 3 can be rewritten as

$$K_{\text{neuro}} \cos \omega_r t = \omega^2_r - K_{\text{mech}}, K_{\text{neuro}} \sin \omega_r t = 2\omega_r \sigma$$
We also know that if sufficient condition for normal breathing is

\[ v = 0, \] 

Otherwise, the earlier assumption of instability border, only if the square root is real, i.e.,

\[ K \]

The above range and the whole solution are meaningful for people with flexible upper airways not to snore. For

\[ v = 0, \] 

as a function of \( v \), the right-hand side (RHS) has a minimum point where \( \omega_r(RHS) = 0, \) which leads to

\[ \omega_r^2 = K_{\text{mech}} - 2\sigma^2, \quad \text{RHS} = 4\sigma^2(K_{\text{mech}} - \sigma^2) \]

We also know that if \( K_{\text{neuro}} \to 0, \) the system goes back to the pure mechanical one, in which \( \omega_r = \sigma > 0. \) So the sufficient condition for normal breathing is

\[ K_{\text{neuro}} < (\text{RHS})_{\text{min}} = 4\sigma^2(K_{\text{mech}} - \sigma^2) \]

which gives the range of damping \( \sigma \) as follows

\[ K_{\text{mech}} - (K_{\text{mech}} - K_{\text{neuro}})^{1/2} < 2\sigma^2 \]

\[ < K_{\text{mech}} + (K_{\text{mech}} - K_{\text{neuro}})^{1/2} \]

The above range and the whole solution are meaningful only if the square root is real, i.e., \( K_{\text{mech}} > K_{\text{neuro}}. \) Otherwise, the earlier assumption of instability border, \( \omega_r = 0, \) is invalid and there will be an unstable mode. This means that one of the necessary conditions for snoring is that the mechanical stiffness should be the dominating force for the maintenance of the upper airway patency. Also, the term \((K_{\text{mech}} - K_{\text{neuro}})^{1/2}\) signifies the margin of stability, which vanishes when \( K_{\text{neuro}} \to K_{\text{mech}}. \) This happens when the upper airway is very flexible, namely, \( K_{\text{mech}} \) is small. It is, therefore, difficult for people with flexible upper airways not to snore.

Another bad respiratory condition is obstructive sleep apnea, which features the complete closure of the passage at some point in the upper airway. The fact that airway stays closed for a considerable period of time means zero angular frequency, \( \omega_r = 0, \) in our analysis. However, an unstable mode also requires a positive rate of growth for small disturbances, \( \omega_i < 0. \) The \( \omega_r = 0 \) means that Eq. 3 is automatically satisfied. Equation 2 then gives

\[ K_{\text{neuro}}e^{\Delta t} + \omega_i^2 - 2\sigma\omega_i + K_{\text{mech}} = 0 \]  

For \( \Delta t = 0, \) we have

\[ \omega_i = \sigma \pm [\sigma^2 - (K_{\text{mech}} + K_{\text{neuro}})]^{1/2} \]

So that if only the combined stiffness is positive, \( K_{\text{mech}} + K_{\text{neuro}} > 0, \) \( \omega_i > 0, \) and no complete collapse will occur. Complete collapse occurs when \( K_{\text{mech}} + K_{\text{neuro}} < 0, \) when the stiffness induced by the muscle activities is not sufficient to sustain the negative pressure of the airflow. The borderline is, of course, \( K_{\text{mech}} + K_{\text{neuro}} = 0. \) We now investigate the effect that the muscle reflex latency \( \Delta t \) has on this borderline condition.

We first notice that \( \omega_i = 0 \) is always a solution to Eq. 7. The other solution is normally positive as \( \Delta t \) increases from 0 until the positive root is merged with \( \omega_i = 0. \) After that, negative roots appear. The condition for the double root for Eq. 7 is

\[ \frac{\partial}{\partial \omega_i} [\omega_i^2 - 2\sigma\omega_i + K_{\text{mech}} + K_{\text{neuro}}e^{\Delta t}]_{\omega_i = 0} = 0 \]

which gives \( \Delta t = 2\sigma/K_{\text{neuro}}. \) As \( \Delta t \to \infty, \) the negative root approaches \( \sigma - (\sigma^2 - K_{\text{mech}})^{1/2}. \)
In summary, we conclude that a mechanical system that is not normally susceptible to total collapse in the absence of latency will become susceptible when the latency exceeds $2\pi/K_{\text{neuro}}$.

Conclusions

First of all, the importance of having a sufficiently rigid airway cannot be overstated. A low tissue (structural) stiffness $K$ will cause initial narrowing of the upper airway, so that the Bernoulli effect lowers the effective stiffness $K_{\text{mech}}$. If $K_{\text{mech}}$ is insufficient to maintain the stability of the airway, neuromuscular functions become crucial. However, these functions are very much reduced during sleep, and the muscle reflex mechanism can have a time delay of several cycles of oscillations experienced during snoring. A delayed restoring force can cause flutter as part of the force is transformed into negative damping. The eigenvalue calculation shows that, depending on the level of the damping coefficient $\sigma$, there is always an eigen mode that is at least marginally unstable. However, there exists a very small margin of damping coefficient for which stable breathing is guaranteed. Sleep apnea occurs when the total upper airway stiffness, $K_{\text{mech}} + K_{\text{neuro}}$ becomes negative, and the airway collapses totally during inspiration. A time delay of $2\pi/K_{\text{neuro}}$ will render neuromuscular function powerless to stop the airway from collapsing.

Should our theoretical model be representative of real human airways, it may provide an avenue for finding new treatments for snoring and sleep apnea. New clinical data are needed to test our hypothesis.

The authors thank Professor A. Guz, whose advice and help were very helpful in keeping the research in touch with clinical practice.

Address for reprint requests and other correspondence: L. Huang, Dept. of Mechanical Engineering, The Hong Kong Polytechnic Univ., Hung Hom, Kowloon, Hong Kong (E-mail: mmlhuang@polyu.edu.hk).

Received 16 June 1997; accepted in final form 1 February 1999.

REFERENCES