

Comment on “Auditory-nerve first-spike latency and auditory absolute threshold: A computer model” [J. Acoust. Soc. Am. 119, 406–417 (2006)] (L)

B. Suresh Krishna^{a)}

*Mahoney Center for Brain and Behavior, Center for Neurobiology and Behavior, Columbia University
College of Physicians and Surgeons, and the New York State Psychiatric Institute, New York, New York 10032*

(Received 22 February 2006; revised 11 May 2006; accepted 18 May 2006)

A recent paper by Meddis [J. Acoust. Soc. Am. 119, 406–417 (2006)] shows that an existing model of the auditory nerve [Meddis and O’Mard, J. Acoust. Soc. Am. 117, 3787–3798 (2005)] is consistent with experimentally-measured first-spike latencies in the auditory nerve [Heil and Neubauer, J. Neurosci. 21, 7404–7415 (2001)]. The paper states that this consistency emerges because in the model, the calcium concentration inside the inner hair cell builds up over long periods of time (up to at least 200 ms) during tone presentation. It further states that integration over long time-scales happens despite the very short time constants (<1 ms) used for the calcium dynamics. This letter demonstrates that these statements are incorrect. It is shown by simulation that calcium concentration inside the hair cell stage of the Meddis model rapidly reaches a steady state within a few milliseconds of a stimulus onset, exactly as expected from the short time-constant in the simple first-order differential equation used to model the calcium concentration. The success of the Meddis model in fitting experimental data actually confirms earlier results [Krishna, J. Comput. Neurosci. 13, 71–91 (2002a)] that show that the experimental data are a natural result of stochasticity in the synaptic events leading up to spike-generation in the auditory nerve; integration over long time scales is not necessary to model the experimental data. © 2006 Acoustical Society of America. [DOI: 10.1121/1.2213569]

PACS number(s): 43.64.Bt [BLM]

Pages: 591–593

Meddis (2006) has recently shown that an existing biophysically detailed model of the auditory nerve (Meddis and O’Mard, 2005) produces first-spike latencies that are qualitatively consistent with experimentally measured first-spike latencies in the auditory nerve (Heil and Neubauer, 2001). The paper asserts (p. 412) that this consistency emerges because “model latencies can be attributed to the accumulation of presynaptic calcium during tone presentation.” The buildup of calcium is ascribed to a mismatch between calcium influx and efflux and is supposed to “operate at least up to tone durations of 200 ms” (p. 412). The paper states that such long time-scale integration over periods up to 200 ms can happen despite the very short time constants (<1 ms) used for the calcium dynamics in the model.

These statements are incorrect, because there is no long-term calcium integration inside the hair cell stage of the Meddis (2006) model. To show this, the calcium concentration inside the inner hair cell stage of the Meddis model in response to a 10 dB sound-pressure level, 200 ms long pure tone was obtained using simulation. In published examples from both the experimental data (Heil and Irvine, 1997) and the Meddis model (Meddis 2006), stimuli at this sound-pressure level and duration fail to elicit a spike after stimulus onset in low spontaneous-rate (SR) fibers on each of 20 presentations; mean first-spike latency was roughly between 10 and 20 ms in high SR fibers.

The inner hair cell potential variation in response to this stimulus is identical for all the parameter variations considered in Meddis (2006); this inner hair cell potential variation serves as the input that drives calcium influx into the hair cell. The inner hair cell potential was obtained using the publicly available Development System for Auditory Modeling (DSAM) of Meddis and O’Mard. All parameter settings were identical to those in Meddis (2006) except that the simulation time step was set to $1 \mu\text{s}$ for improved accuracy, instead of the $10 \mu\text{s}$ in Meddis (2006); the output inner hair cell potential from DSAM was essentially unchanged by this change in time step. The remaining stages between the inner hair cell potential and calcium concentration were simulated in MATLAB using a $1 \mu\text{s}$ time-step: this includes Eqs. (A7)–(A10) in Meddis (2006).

Figure 1(a) plots the inner hair cell potential obtained using DSAM in response to a 200 ms long pure tone (at 10 dB sound pressure level and 4 kHz frequency) shaped by a cosine-squared ramp with 1.7 ms rise/fall time at onset and offset, respectively. According to Meddis (2006, p. 410), such a stimulus fails to elicit a spike in low SR fibers possibly because “the time taken to accumulate enough presynaptic calcium to initiate a spike is greater than the duration of the signal,” clearly implying that calcium is building up over the entire duration of the tone, but at too slow a rate to elicit a spike. Figures 1(c) and 1(e) plot the calcium concentration inside the inner hair cell with parameters set for low SR fibers exactly as in Meddis (2006). Calcium concentration rises rapidly to reach 99% of its peak deviation from baseline within about 4.97 ms after stimulus onset. Since the inner

^{a)}Electronic mail: ssk2031@columbia.edu

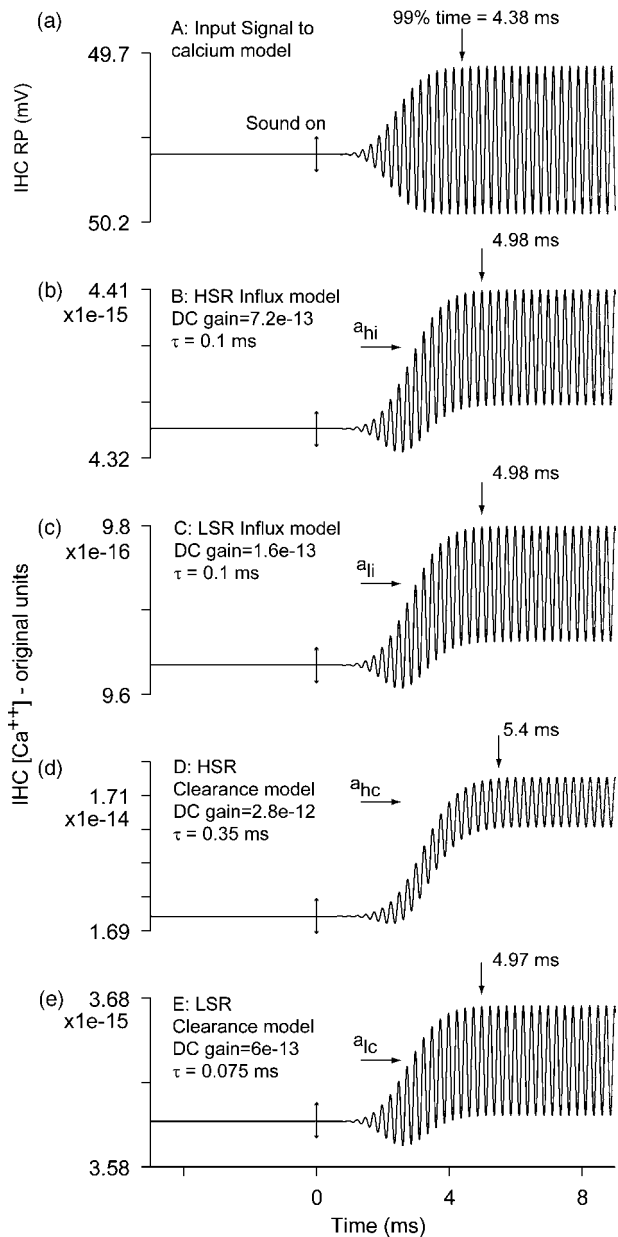


FIG. 1. Calcium concentration in the Meddis inner hair cell reaches steady state within a few milliseconds of sound onset: there is no long-term buildup of calcium. (a) Plot of inner hair cell receptor potential (mV) as a function of time (ms) in response to a 4 kHz tone (10 dB sound pressure level, 1.7 ms cosine-squared rise-fall times). Small double-headed arrow at time=0 ms indicates sound onset. Also indicated is the time required for the potential to reach 99% of its peak excursion; in this case, it is 4.38 ms. (b)–(e) Plots of calcium ion concentration [units identical to Meddis (2006)] in the inner hair cell as a function of time (ms) with parameters set to those of the high SR influx model (b), the low SR influx model (c), the high SR clearance model (d), and the low SR clearance model (e), respectively. The rightward arrows labeled a_{hi} , a_{li} , a_{hc} , and a_{lc} show the steady-state dc component of calcium concentration, obtained as the mean calcium concentration from 100 to 150 ms after sound onset, for each model. The dc gain (obtained as $\tau_{Ca} * G_{Ca}^{max}$) and time constant of calcium clearance (τ_{Ca} : shown as τ) of the models are indicated in each graph. Sound onset and time for ordinate value to reach 99% of its peak excursion are shown as in (a).

hair cell potential reaches 99% of its peak deviation from baseline about 4.38 ms after stimulus onset, it is clear that the added lag due to the dynamics of calcium influx and clearance is extremely small. Importantly, the calcium concentration does not steadily increase over the duration of the tone, as implied by Meddis (2006) on the basis that calcium

dynamics show long-term integration. The results for high SR fibers are similar. Since the time constants in the influx model are identical for low, medium, and high SR fibers in Meddis (2006), the time for calcium concentration in the high SR case to reach 99% of peak deviation [Fig. 1(b)] is identical to that for low SR fibers [Fig. 1(c)]. In the clearance model, the time constant (τ_{Ca}) of clearance for high SR fibers is slightly longer, and therefore the time to reach 99% of peak deviation is also slightly longer by about 0.4 ms [Figs. 1(d) and 1(e)]. Importantly, in contrast to the similarity in rise times, there is a large difference in both transient and average steady-state concentrations between high and low SR fibers; this is true for both influx and clearance models. This difference in concentrations is expected from the analytical expression for the transfer gain of a first-order low-pass filter. For frequency components that are small compared to $(1/\tau_{Ca})$, the gain relating calcium concentration to the input (third power of proportion of open channels multiplied by driving potential, after absorbing the G_{max}^{Ca} term into the gain) in the Meddis model [Eq. (A7)] is approximately equal to $(G_{max}^{Ca} * \tau_{Ca})$; the expression is exactly correct at a frequency of 0 (dc). For the influx model, only G_{max}^{Ca} is changed (by a factor of 4.5) to model the differences in calcium dynamics between high and low SR fibers; therefore the only change in model output is a simple scaling that is independent of frequency. The model calcium concentration for high SR fibers is simply 4.5 times the model calcium concentration for low SR fibers. For the clearance model, since τ_{Ca} is varied between high and low SR fibers, the gain is frequency dependent; however, at dc, the gain ratio is 4.67 and therefore the average steady-state calcium concentration for high SR fibers is 4.67 times the average steady-state concentration for low SR fibers. But neither fiber type shows any sign of long-term integration in the calcium concentration profile.

The Meddis (2006) model fits experimental latency data not because of long-term calcium integration, but because of its stochastic spike-generation process. Ignoring refractory effects from spikes prior to stimulus onset, the Meddis model generates a spike whenever any of q available quanta is released, with each of the q quanta having a fixed small stimulus-dependent probability p of generating a spike in a single simulation epoch. If the mean rate of quantal release is higher, then the mean first-spike latency will be lower. For small values of p (p can be made as small as necessary by decreasing the simulation time step dt), and for purposes of calculating first-spike latency, the spike train from the Meddis model's spiking algorithm can be approximated by a Poisson process with a time varying firing rate roughly equal to qp/dt . If this spike train fires at a mean rate of 10 spikes/s, a reasonable value for firing rates of low-SR fibers to near-threshold stimuli, the expected mean first-spike latency is 100 ms. Thus, this approximate calculation shows that the mean first-spike latencies are essentially consistent with the firing rates obtained from auditory nerve fibers. Further, this reasoning also shows why changes in SR in the Meddis model lead to differences in first-spike latency. The main effect of the parameter changes that are used to model differences in SR in the Meddis (2006) model is an approxi-

mately 4.5-fold difference in gain. This difference in gain is the primary contributor to the observed latency effects; the changes in rise time (i.e., “integration time”) between fibers of different SR are minor (zero for the influx model, and less than 0.5 ms for the clearance model: see Fig. 1). When the gain is higher, so is the driven spike rate and the mean first-spike latency is therefore smaller.

The above-presented argument is a summary of the one in Krishna (2002a), where it was shown that the experimental data can be captured at a quantitative level by a minimal model that relies upon stochastic spiking at a rate proportional to an *almost* instantaneous function of sound pressure; no long-term integration of stimulus magnitude is required to capture the data. This minimal model was a simplified version of current standard models of the auditory periphery, including the Meddis (2006) model. Krishna (2002a, p. 86) implied that all such standard models could be expected to capture the data successfully; indeed, this is exactly what Meddis (2006) confirms. The above-noted argument also enables a long-due clarification of the physiological basis for the Poisson model used in Krishna (2002a). Krishna (2002a) did not explicitly localize the physiological basis for his spike-generation process. Specifically, despite what the regrettably confusing phrasing in one section might appear to possibly imply, the paper did not intend to suggest that the stochasticity in spike generation arises postsynaptically (see criticism in Heil, 2004). The most likely physiological basis for the Poisson model is the approximately Poisson nature of the vesicle release process; if each random vesicle release leads to a reliably timed spike as has been suggested in the physiological literature (Siegel, 1992), then the spike train will also possess Poisson properties (as explicitly stated in Krishna 2002b, p. 216). Available extracellular data strongly support the use of a Poisson process (modified by a refractory period) to model auditory nerve spike-train data (Johnson, 1978). It is important to emphasize that just like the model in Krishna (2002a), the Meddis model also contains what Meddis (2006, p. 407) calls an “explicit integration module;” this integration module resides not in the calcium dynamics, but in the stochastic vesicle release process. This is because models that characterize the analog to discrete-event conversion stage of auditory transduction (including the processes of vesicle release and post-synaptic spiking) as a point process can equivalently be interpreted as incorporating a stage where the conditional intensity function of the point process is integrated until the integral crosses an exponentially distributed random threshold. This fact was stated for the Poisson case in Krishna (2002a); the general version of this theorem can be found in Brown *et al.* (2002).

Krishna (2002a) did not imply that calcium dynamics have no role to play in determining first-spike latency, as stated in Meddis (2006). The specific objection in Krishna (2002a) was to the idea that the latency data were best modeled using a calcium integrator that, to paraphrase Heil and Neubauer (2001, p. 7414), can be “conceptualized using the metaphor of a barrel that has an inflow component that is directly proportional to stimulus peak-pressure and an inflow/outflow component that is constant; the first spike is triggered when the amount of fluid in the barrel reaches a

critical value.” It is not at all obvious how such a “barrel” model would be consistent with the known ability of auditory-nerve fibers to follow stimulus variations up to several kilohertz. It is important to note that Heil and Neubauer have since explicitly disclaimed this “barrel” metaphor and state that “we do not assume the existence of a response which is directly proportional to the continuously increasing integral of $P(t)$ or of $P(t)$ raised to some power and which could be measured during a given trial;” here $P(t)$ refers to the peak pressure of the stimulus (Neubauer and Heil, 2004, p. 438). Instead, they now favor a stochastic interpretation (Heil and Neubauer, 2003; Neubauer and Heil, 2004) that is somewhat different from the one discussed earlier. This explanation, in apparent agreement with Krishna (2002a), also relies partially upon the fact that long first-spike latencies will emerge from low driven spike rates. However, from their cursory description, it is unclear how their empirical law governing spike timings (and hence the spike rate) can be derived from their stochastic formulation of calcium dynamics. Further, while the details are unclear, this stochastic formulation still appears to retain a similar problematic dependence upon a long-term running average of stimulus pressure. It remains to be demonstrated that their new proposal is consistent with other aspects of the data like the variability of first-spike latency and the ability of auditory-nerve fiber responses to follow high stimulus frequencies. In any event, the “widely-held objection” (Meddis 2006, p. 407) to the idea that short-term submillisecond dynamics in the auditory periphery can be used to create long-term temporal integration over periods up to 100 ms or more is well-founded and remains entirely valid.

The author thanks Laurel Carney for reading and commenting on the first draft.

- Brown, E. N., Barbieri, R., Ventura, V., Kass, R. E., and Frank, L. M. (2002). “The time-rescaling theorem and its application to neural spike train data analysis,” *Neural Comput.* **14**, 325–346.
- Heil, P. (2004). “First-spike latency of auditory neurons revisited,” *Curr. Opin. Neurobiol.* **14**, 461–467.
- Heil, P., and Irvine, D. R. (1997). “First-spike timing of auditory-nerve fibers and comparison with auditory cortex,” *J. Neurophysiol.* **78**, 2438–2454.
- Heil, P., and Neubauer, H. (2001). “Temporal integration of sound pressure determines thresholds of auditory-nerve fibers,” *J. Neurosci.* **21**, 7404–7415.
- Heil, P., and Neubauer, H. (2003). “A unifying basis of auditory thresholds based on temporal summation,” *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6151–6156.
- Johnson, D. H. (1978). “The relationship of post-stimulus time and interval histograms to the timing characteristics of spike trains,” *Biophys. J.* **22**, 413–430.
- Krishna, B. S. (2002a). “A unified mechanism for spontaneous-rate and first-spike timing in the auditory nerve,” *J. Comput. Neurosci.* **13**, 71–91.
- Krishna, B. S. (2002b). “Temporal response properties of single neurons in the inferior colliculus of the anesthetized gerbil,” PhD dissertation, Center for Neural Science, New York University.
- Meddis, R. (2006). “Auditory-nerve first-spike latency and auditory absolute threshold: A computer model,” *J. Acoust. Soc. Am.* **119**, 406–417.
- Meddis, R., and O’Mard, L. P. (2005). “A computer model of the auditory-nerve response to forward-masking stimuli,” *J. Acoust. Soc. Am.* **117**, 3787–3798.
- Neubauer, H., and Heil, P. (2004). “Towards a unifying basis of auditory thresholds: The effects of hearing loss on temporal integration reconsidered,” *J. Assoc. Res. Otolaryngol.* **5**, 436–458.
- Siegel, J. H. (1992). “Spontaneous synaptic potentials from afferent terminals in the guinea pig cochlea,” *Hear. Res.* **59**, 85–92.