

**Modelling envelope and temporal fine structure components of frequency-following responses in rat inferior colliculus**

WANG Qian and LI Liang

Citation: [SCIENCE CHINA Technological Sciences](#) **60**, 966 (2017 ); doi: 10.1007/s11431-016-9044-5View online: <http://engine.scichina.com/doi/10.1007/s11431-016-9044-5>View Table of Contents: <http://engine.scichina.com/publisher/scp/journal/SCTS/60/7>Published by the [Science China Press](#)**Articles you may be interested in**[FINE STRUCTURE AND MORPHOGENESIS OF GOAT POXVIRUS ENVELOPE](#)Chinese Science Bulletin **31**, 341 (1986);[FINE STRUCTURE AND MORPHOGENESIS OF GOAT POXVIRUS ENVELOPE](#)Chinese Science Bulletin **32**, 341 (1987);[Role of frequency band integration in sharpening frequency tunings of the inferior colliculus neurons in the big brown bat, \*Eptesicus fuscus\*](#)Chinese Science Bulletin **49**, 1026 (2004);[FINE STRUCTURE OF INFECTIOUS BOVINE RHINOTRACHEITIS VIRUS \(IBRV\) ENVELOPE](#)Chinese Science Bulletin **30**, 114 (1985);[Effects of forward masking on the responses of the inferior collicular neurons in the big brown bats, \*Eptesicus fuscus\*](#)Chinese Science Bulletin **48**, 1748 (2003);

WANG Qian<sup>1,2</sup> & LI Liang<sup>1,3,4\*</sup>

*School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University,  
Beijing 100871, China;*

<sup>2</sup>*Beijing Key Laboratory of Epilepsy, Epilepsy Center, Department of Functional Neurosurgery, Sanbo Brain Hospital, Capital Medical  
University, Beijing 100093, China;*

<sup>3</sup>*Speech and Hearing Research Center, Key Laboratory on Machine Perception (Ministry of Education), Peking University,  
Beijing 100871, China;*

<sup>4</sup>*Beijing Institute for Brain Disorders, Beijing 100081, China*

0

In studies of auditory perception, a dichotomy between envelope and temporal fine structure (TFS) has been emphasized. It has been

of neuron populations to instantaneous waveforms of low-to-middle-frequency acoustic stimuli [12–20]. Previous studies had shown that FFRs can efficiently convey both TFS signals [12,14,21–25] and envelope signals (also called envelope-following response) [25–31]. As FFRs have been thought to be mainly originated from the auditory mid-brain (i.e. inferior colliculus, IC) [12,13,18,32,33], our recent studies have also further suggested a functional dichotomy of the envelope component of FFR ( $FFR_{Env}$ ) and the temporal fine structure component of FFR ( $FFR_{TFS}$ ) in the rat IC [25].

Computational models for the generation of FFR are critical for understanding the properties of FFR, based on the neurophysiological parameters along the ascending pathway. Up to date, there are two existing FFR models. One has been established for human scalp FFRs [34] by convoluting click-evoked waveforms with the instantaneous discharge rates based on nonlinear cochlear processing [35]. Dau's FFR model ( $FFR$ ) predicts well for the frequency dependent latency and intensity effect, but fails to simulate binaural response properties raised from brainstem nuclei.

The other one has been established for extracellular field potentials in owl's medial superior olivary nuclei (MSO) [36] by convolving Poisson processed spike trains with a model spike waveform. Although Kuokkanen et al.'s [36] FFR model solved the interaural processing pattern by linear summation of ipsilateral and contralateral input, some properties such as nonlinear cochlear processing or cellular properties are ignored. Thus, neither of these two models can separate  $FFR_{Env}$  and  $FFR_{TFS}$ .

Can different cellular properties of IC neurons be used to distinguish  $FFR_{Env}$  from  $FFR_{TFS}$ ? The temporal processing properties of IC neurons can be described by modulation transfer functions (MTFs), which include several types [37]. A recent computational model has shown that activations of IC neurons with different types of MTFs can be effectively tuned to signal fluctuations [38]. More specifically, IC neurons with bandpass MTFs tune to slowly fluctuated amplitude modulations while neurons with band-reject or low-pass MTFs tune to quickly fluctuated vowel formants. These findings imply that the MTF types of IC neurons can be used to establish an IC FFR model which effectively separates  $FFR_{Env}$  and  $FFR_{TFS}$ .

On the purpose to develop an appropriate IC FFR model, we combined the advantages of these two existing FFR models [34,36], and simultaneously added a construction of MTF types IC neurons to simulate the separated envelope and TFS information processing [38]. To evaluate the validity of the current model, narrowband noise evoked intracranial FFRs were recorded in rat IC. Different model parameters were tested, correlations between simulated FFRs and experimental FFRs were calculated, and the parameters with highest prediction were chosen to build an ideal FFR model.

Intracranial IC FFRs were measured in 12 young-adult male Sprague-Dawley rats (age 10–12 weeks, weight 280–350 g). All rats were treated in accordance with both *the Guidelines of the Beijing Laboratory Animal Center* and *the Policies on the Use of Animal and Humans in Neuroscience Research* approved by the Society for Neuroscience (2006).

During the recording, rats were anesthetized with 10% chloral hydrate (400 mg/kg, intraperitoneal) and the state of anesthesia was maintained throughout the experiment by supplemental injection of the same anesthetic. Stainless steel recording electrodes (10–20 k $\Omega$ ), which were insulated by a silicon tube (0.3 mm in diameter) except at the 0.25 mm diameter tip [13,15,25], were aimed at the central nucleus of the IC bilaterally. Based on the stereotaxic coordinates of ref. [39] and referenced to *Bregma*, the IC coordinates were: AP, –8.8 mm; ML,  $\pm$ 1.5 mm; DV, –4.5–5.0 mm.

Gaussian wideband noises (10-kHz sampling rate and 16-bit amplitude quantization) were generated and filtered by a 512-point digital filter with a center frequency of 2000 Hz and a bandwidth of 0.466 octaves (ranged from 1680 to 2320 Hz) with MATLAB (MathWorks, Natick, MA, USA). The stimulus duration was 700 ms with 10-ms linear onset/offset ramps, and the (offset-onset) inter-stimulus interval was 100 ms. All sound waves were re-processed by a TDT System II (Tucker-Davis Technologies, FL, USA), and presented through two ED1 earphones. Two 12-cm TDT sound-delivery rubber tubes were connected to the ED1 earphones and inserted into each of the rat's ear canals for sound delivery. All narrow-band noises were calibrated using a Larson Davis Audiometer Calibration and Electroacoustic Testing System (AUDITM and System 824, Larson Davis, USA). The level of all signals was 72 dB SPL for each earphone.

Evoked neural potentials were recorded in a sound-attenuating chamber, amplified 1000 times by TDT DB4 amplifier, filtered through a 100–10000 Hz band-pass filter (with a 50-Hz notch), and averaged 100 times per stimulus condition. Online recordings were processed with TDT Biosig software, digitized at 16 kHz, and stored on disk for off-line analyses. Stimuli were delivered to rats monaurally (either ipsilaterally or contralaterally).

When all recordings were completed, rats were euthanized with an overdose of chloral hydrate. Lesion marks were made via the recording electrodes by an anodal DC current (500  $\mu$ A for 10 s). The brains were stored in 10% formalin with 30% sucrose, and then sectioned at 55  $\mu$ m in the frontal plane in a cryostat (–20°C).

Theoretically, a steady-state Gaussian narrowband noise with a center frequency of  $c$  Hz and a bandwidth of  $b$  Hz has a TFS energy around  $c$  Hz and an envelope energy within

the frequency range between 0 and  $b$  Hz [40]. Thus, for a narrowband noise with bandwidth 640Hz, the envelope spectrum were ranged from 0 to 640Hz, while the TFS spectrum were ranged from 1680 to 2320Hz. In that case, low-pass and high-pass FIR filters with 1000Hz cut-off frequency were designed by the MATLAB `fr2` function (512th-order) to extract the  $\text{FFR}_{\text{TFS}}$  and  $\text{FFR}_{\text{Env}}$  components from the original potential, respectively.

As shown in Figure 1, the current FFR model in the present study mainly contented 5 steps. (1) The incoming stimulus was firstly processed through the newest computational auditory nerve (AN) [41]. This phenomenological AN model included several key nonlinearities, rate saturation, adaptation, and synchrony capture. (2) The second stage used two different rMTF functions (band-pass, BP; band-reject, BR) to separate the envelope and TFS responses [38]. (3) Then, the outputs from the second stage, instantaneous discharge rate along frequency channels, were produced by the fast-Poisson process [42] to generate spike trains. (4) The spike train in each frequency channel was then convolved with a model spike waveform  $k$ , which was approximated with a Gabor function [36].

$$k(t) \propto \exp\left(-\frac{t^2}{2\rho^2}\cos(2\pi f_g t + \phi)\right), \quad (1)$$

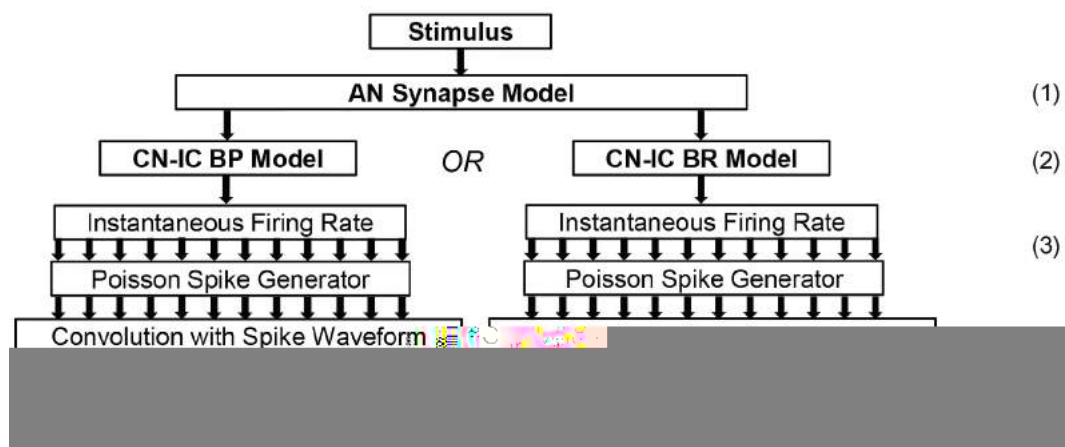
where width  $\rho = 0.09$ ms, primary phase  $\phi = 0.8$ rad, oscillation frequency  $f_g = 3.9$ kHz. The absolute peak amplitude was set to  $100 \mu\text{V}$ . (5) The fourth stage mentioned above was generated independently by 100 times. The simulated FFR potentials were generated by summing the output of each CF channels, with 100 repetitions.

Narrowband noise with sound level of 72dB SPL, center frequency of 2000Hz, and bandwidth of 640Hz, which was used to evoke the experimental FFR in the present study was entered into the model. The fixed model parameters were set as follows: 100-kHz sampling rate; best-frequency resolution of 50Hz (from 125 to 5000Hz).

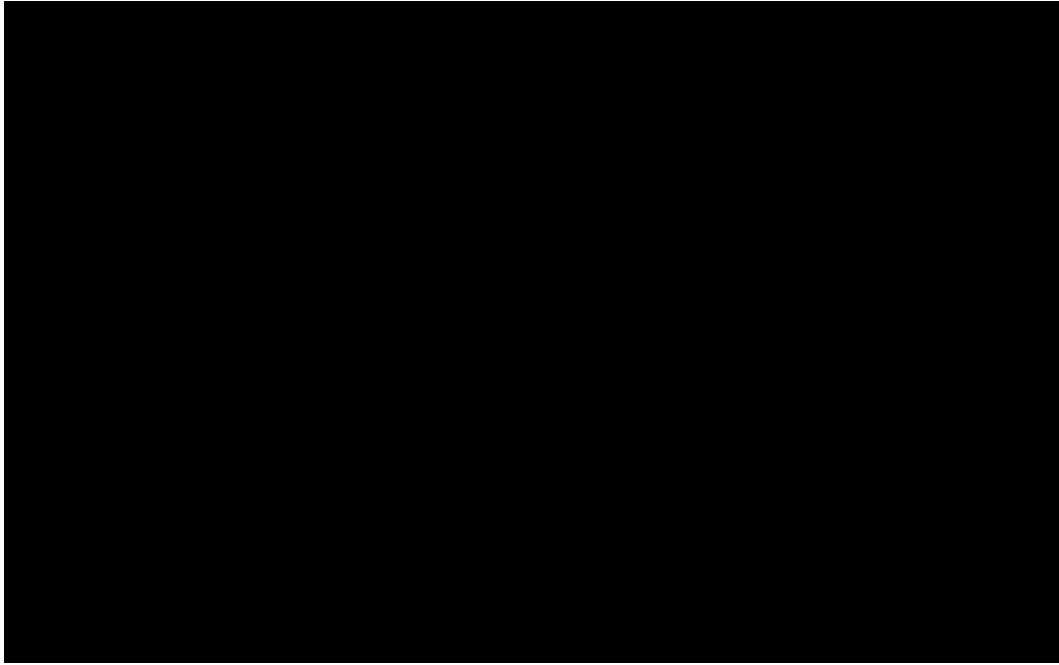
Three model parameters [AN spontaneous rates: 0.1 (low spontaneous rate, LSR), 4 (middle spontaneous rate, MSR), and 100 (high spontaneous rate, HSR) spikes/s; best modulation frequencies (BMFs): 16, 32, 64, 128, and 256Hz; center frequencies (CFs): 0.125–1.5, 1.5–5 kHz] had been tested and model FFR of each parameter combinations. Prediction indices (PIs) of  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$  were calculated as Pearson correlation coefficients between real FFR waveforms and simulated FFR waveforms, respectively. The optimized parameter combinations were conducted into the ideal FFR model.

Statistical analyses were performed with MATLAB (MathWorks, Natick, MA, USA). Pairwise  $t$ -tests were conducted and the null-hypothesis rejection level was set at 0.05.

Based on our previous study [25], an effective IC FFR model should be able to separate  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$ . In the current model, spike trains output within all CF channels (from 125 to 5000Hz) from CN-IC BP model showed relative long periodicity (Figure 2(a)) while those from CN-IC BR showed relative short periodicity (Figure 3(a)). In that case, the simulated  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$  (Figures 2(b) and 3(b)) differed in their spectra (Figures 2(c) and 3(c)): the spectra of simulated  $\text{FFR}_{\text{Env}}$  distributed mainly below 640Hz (same frequency range of the acoustic envelope), while the spectra



Schematic of the model structure of the inferior colliculus (IC) originated  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$ , respectively. (1) An existing auditory peripheral model (auditory nerve model [41]) was used to provide inputs to higher neural center (frequency channels centered from 125 to 5000Hz,  $f = 50$ Hz). (2) Then, at CN-IC level, band-pass (BP) and band-reject (BP) models [38] were used to separate the envelope and TFS information. (3) The instantaneous firing rate of each frequency channel was used to generate spike trains by fast-Poisson process [42]. (4) The spike trains were then convolved with a model spike waveform to get the field potentials. (5) Finally, the simulated FFRs were summed cross all the potential in frequency channels, with 100 independent repetitions. IC, inferior colliculus; IHC, inner hair cell; AN, auditory nerve; CN, cochlear nucleus; BP, band-pass; BR, band-rejected.



The raw  $\text{FFR}_{\text{Env}}$  model outputs under different best modulation frequency (BMF) conditions (high spontaneous rates of modeled AN fibers). (a) Spike trains in all frequency channels (from 125 to 5000Hz); (b) waveforms of model  $\text{FFR}_{\text{Env}}$ ; (c) spectra of model  $\text{FFR}_{\text{Env}}$ . AN, auditory nerve.

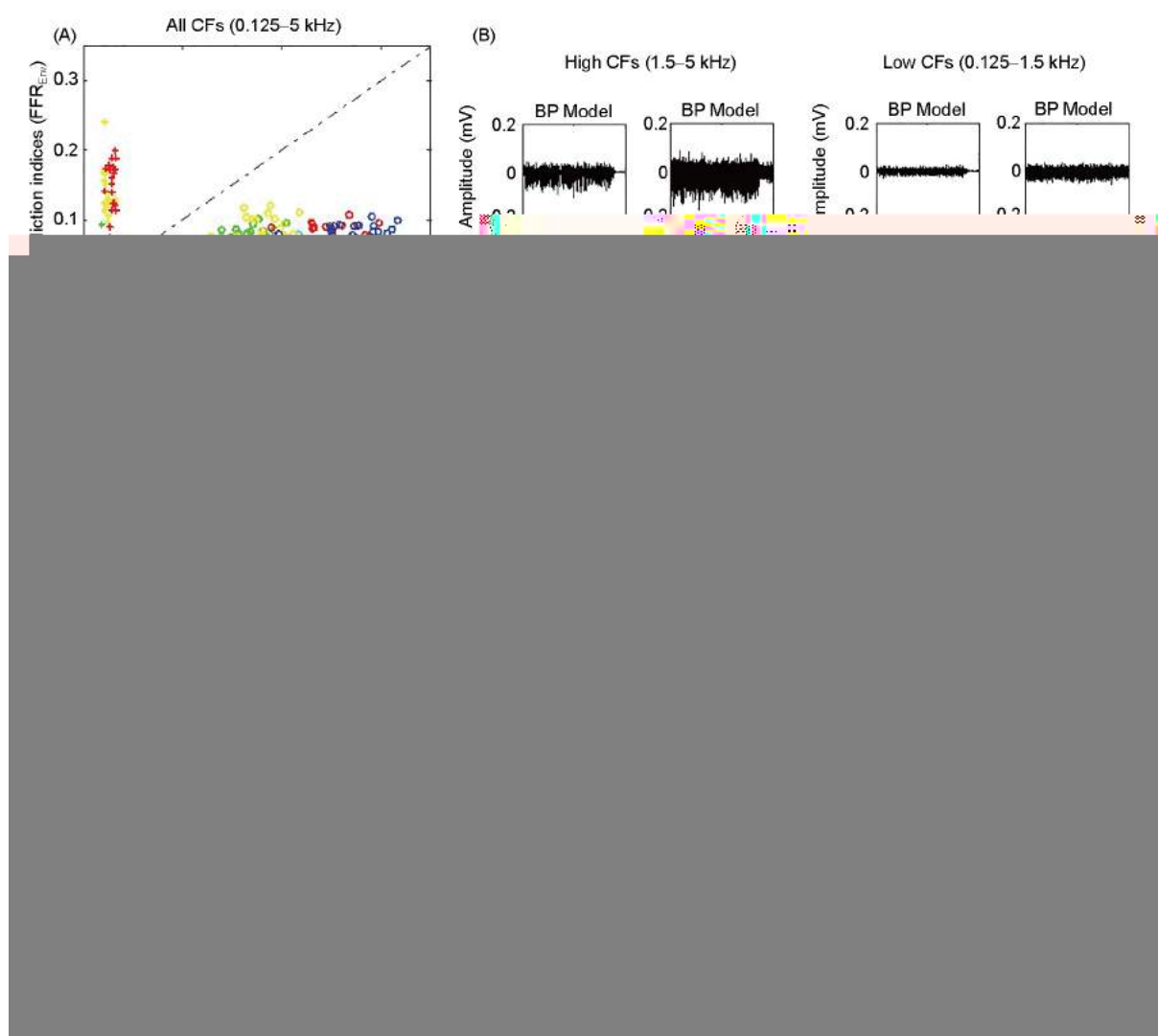


The raw  $\text{FFR}_{\text{TFS}}$  model outputs under different best modulation frequency (BMF) conditions (high spontaneous rates of modeled AN fibers). (a) Spike trains in all frequency channels (from 125 to 5000Hz); (b) waveforms of model  $\text{FFR}_{\text{TFS}}$ ; (c) spectra of model  $\text{FFR}_{\text{TFS}}$ . AN, auditory nerve.

simulated  $\text{FFR}_{\text{TFS}}$  distributed around 2000Hz (same frequency range of the acoustic TFS).

As shown in Figures 2 and 3, the simulated FFRs of different BMFs differed both in their waveforms and spectra. In order to estimate the model predictive validity, PIs of  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$  were calculated as Pearson correlation coefficients between real FFR waveforms and simulated FFR waveforms, respectively. Figure 4(A) presents how well the

simulated  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$  could effectively imitate the real  $\text{FFR}_{\text{Env}}$  and  $\text{FFR}_{\text{TFS}}$  along different BMFs. Ideally, the PIs of  $\text{FFR}_{\text{TFS}}$  of model TFS should be larger than those of model envelope, while the PIs of  $\text{FFR}_{\text{Env}}$  of model envelope should be larger than those of model TFS. In the other word, in Figure 4(A), the PIs of  $\text{FFR}_{\text{TFS}}$  (circles in Figure 4(A)) should be distributed below the diagonal and the PIs of  $\text{FFR}_{\text{Env}}$  should be distributed above the diagonal (crosses in



Illustrations of the model  $FFR_{Env}$  and  $FFR_{TFS}$  generated by low (0.125–1.5 kHz) and high (1.5–5 kHz) frequency channels only. (A) Comparison of prediction indices (PIs) of  $FFR_{Env}$  or  $FFR_{TFS}$ . For each dot, the distance from the diagonal represented the discrimination degree between  $FFR_{Env}$  or  $FFR_{TFS}$ . (B) Examples of model  $FFR_{Env}$  or  $FFR_{TFS}$  generated by low and high CFs respectively. The model  $FFR$  generated by low CFs gave quite low signal-to-noise ratio within  $FFR$  spectrum. (C) Comparisons of model predictions to experimental  $FFR_{Env}$  and  $FFR_{TFS}$  with different model parameters. The low CFs generated model  $FFRs$  gave a weak contributions to the experimental data. BMF, best modulation frequency; CFs, frequency channels; M, model.

Figure 4(A)). Furthermore, the distance from the diagonal represents the discrimination degree between  $FFR_{Env}$  and  $FFR_{TFS}$ . The top panel of Figure 4(A) shows that the current  $FFR$  model can effectively simulate both  $FFR_{Env}$  and  $FFR_{TFS}$  and effectively separated the two components.

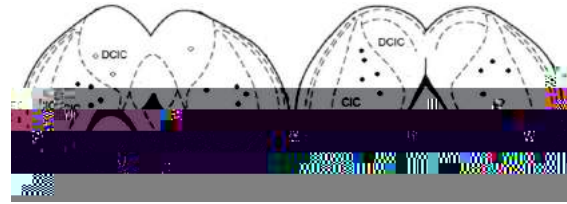
In order to choose the optimized parameters, two parameters of the current model (AN spontaneous rate: LSR, MSR, and HSR; BMF: 16, 32, 64, 128, 256 Hz) were modulated. As the previous studies [34] have suggested that high-frequency channels (from 1.5 to 10 kHz) contribute more to  $FFR$ , two different CF channels (low CFs: 0.125–1.5 kHz; high CFs: 1.5–5 kHz) were also modulated for the comparison. The PIs of  $FFR_{Env}$  and  $FFR_{TFS}$  under each parameter combination were calculated for further analyses. As shown in Figure 4(A), PIs of model  $FFR_{Env}$  and  $FFR_{TFS}$  obtained

with high frequency channels were similar to those obtained with all CF channels (0.125–5 kHz), which significantly contribute to experimentally obtained  $FFR_{Env}$  and  $FFR_{TFS}$ . However, PIs of model  $FFR_{Env}$  and  $FFR_{TFS}$  obtained with low frequency channels were poor (bottom panel of Figure 4(A)). An example of simulated  $FFR_{Env}$  and  $FFR_{TFS}$  (HSR, BMF = 128 Hz) was that either  $FFR_{Env}$  or  $FFR_{TFS}$  obtained by low frequency channels were lower in the absolute amplitude of signal-to-noise ratio in spectra (Figure 4(B)). These results demonstrate that both low-frequency envelope and high-frequency TFS arise from the fine phase-locking mechanisms in the high-frequency channels.

According to the PIs values of  $FFR_{Env}$  and  $FFR_{TFS}$ , we developed the monaural ideal  $FFR_{Env}$  and  $FFR_{TFS}$  model by averaging the outputs selections of optimized parameter

combinations. As shown in Figure 4(C), the optimized selections include two for  $FFR_{TFS}$  (HSR, BMF = 128 and 256Hz) and four for  $FFR_{Env}$  (HSR and MSR  $\times$  BMF = 64 and 128Hz), only obtained with high CFs. Compared to PIs of raw model, PIs of ideal simulated  $FFR_{TFS}$  and  $FFR_{Env}$  were significantly higher under both ipsilaterally and contralaterally stimulated conditions (Figure 5) (paired  $t$ -tests, all  $p < 0.05$ ). On the other hand, PIs of ideal simulated  $FFR_{TFS}$  and  $FFR_{Env}$  were significantly higher than PIs of raw model under both ipsilaterally and contralaterally stimulated conditions (Figure 5) (paired  $t$ -tests, all  $p < 0.05$ ).

To be noted, all the experimental data used was from the central nucleus of IC in 21 out of the 24 recording sites, according to the histological examination (Figure 6).

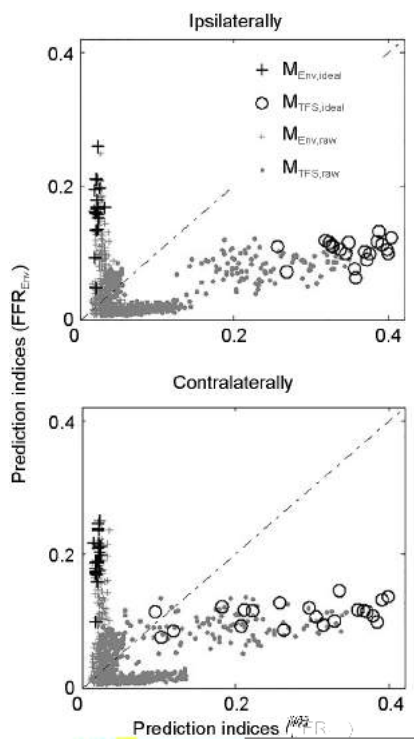


Histological results of recording electrodes in 12 rats. Electrodes were precisely located within the central nucleus of the inferior colliculus (IC) in 21 of the 24 penetrations (filled circles and the star). Note that two electrodes were inserted per animal, one on each side of the IC.

[25], we used  $\epsilon$  TFS . (Gq)  $\epsilon$   $\epsilon$

P) # 7 SR

The present study investigates the roles of both auditory periphery processing (parameters: spontaneous rates and CFs) and auditory brainstem processing (parameters: BMF) for the formation of  $FFR_{Env}$  and  $FFR_{TFS}$ , respectively. According to our previous intracranial FFR recording studies in the rat IC



Outputs of the monaural ideal  $FFR_{Env}$  and  $FFR_{TFS}$  model. Four optimized parameter combination (HSR, MSR  $\times$  BMF = 64, 128Hz) were chosen for the ideal  $FFR_{Env}$  model while two (HSR, BMF = 128, 256Hz) were chosen for the ideal  $FFR_{TFS}$  model. For each data, the distance from the diagonal represented the discrimination degree between  $FFR_{Env}$  and  $FFR_{TFS}$ . Compared to the raw model data, the ideal FFR model predictions to ipsilateral  $FFR_{TFS}$  and contralateral  $FFR_{Env}$  was significantly improved.

from his previous study [44] in order to predict the scalped recorded FFRs. While in the current study, the unitary response function is calculated as a model spike waveform  $k$ , which is approximated with a Gabor function [36,45] in order to predict the intracranial recorded FFRs. Comparing with Kuokkanen's FFR model [36], the advantages of the new developed model is the ability to separate neural TFS and envelope components by a phenomenological modeling approach. Furthermore, because the Kuokkanen's FFR model [36] is established for birds' MSO, the new developed model which is established for rats' IC might be more suitable to interpret the response patterns of FFRs in mammals. The gap between prediction of scalp FFRs and intracranial FFRs should be fixed in the future investigations.

In conclusion, an effective FFR model which explains the formation of  $FFR_{Env}$  and  $FFR_{TFS}$  is developed. This model can be further applied to different forms of hearing disorder to understand what stages the impairments of FFR generation happen.

*We thank Laurel Carney for discussion and for critical reading of an earlier version of this paper. This work was supported by the National Natural Science Foundation of China (Grant No. 31470987), the National Basic Research Development Program of China (Grant No. 2015CB351800), "985" grants from Peking University for Physiological Psychology and China Postdoctoral Science Foundation (Grant No. 2016M601066).*

- 1 Moore B C J. The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Oto*, 2008, 9: 399–406
- 2 Rosen S. Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos Trans R Soc B-Biol Sci*, 1992, 336: 367–373
- 3 Smith Z M, Delgutte B, Oxenham A J. Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 2002, 416: 87–90

J F Fousti ! FFr , f  
 # ]d 3 fJ, ofso ,nd ~ ic ure , fJ# P Y-1" T = U.a i iq



- (*Tyto alba*). *J Neurophysiol*, 2010, 104: 2274–2290
- 37 Langner G, Schreiner C E, Bielel U W. Functional implications of frequency and periodicity coding in auditory midbrain. In: Palmer A R, Rees A, Summerfeld A Q, et al., eds. *Psychophysical and Physiological Advances in Hearing*. London: Whurr Publ. Ltd., 1998. 277–285
- 38 Carney L H, Li T, McDonough J M. Speech coding in the brain: Representation of vowel formants by midbrain neurons tuned to sound fluctuations. *eNeuro*, 2015, 2
- 39 Paxinos G, Watson C, Carrive P, et al. *Chemoarchitectonic Atlas of the Rat Brain*. Atlanta: Elsevier, 2009
- 40 Longtin A, Middleton J W, Cieniak J, et al. Neural dynamics of envelope coding. *Math Biosci*, 2008, 214: 87–99
- 41 Zilany M S A, Bruce I C, Carney L H. Updated parameters and expanded simulation options for a model of the auditory periphery. *J Acoust Soc Am*, 2014, 135: 283–286
- 42 Jackson. The SGfast Mex Function. <https://www.urmc.rochester.edu/MediaLibraries/URMCMedia/labs/carney-lab/documents/articles/Jackson-SGfast-2003.pdf>, 2003
- 43 Goldstein Jr. M H, Kiang N Y S. Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. *J Acoust Soc Am*, 1958, 30: 107–114
- 44 Dau T, Wegner O, Mellert V, et al. Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J Acoustical Soc Am*, 2000, 107: 1530–1540
- 45 Gold C, Henze D A, Koch C, et al. On the origin of the extracellular action potential waveform: a modeling study. *J Neurophysiol*, 2006, 95: 3113–3128