frontiers in
NEURAL CIRCUITS

ORIGINAL RESEARCH ARTICLE
published: 21 October 2013
doi: 10.3389/fncir.2013.00166

Neural representation in the auditory midbrain of the envelope of vocalizations based on a peripheral ear model

Thilo Rode[1]*, Tanja Hartmann[1], Peter Hubka[1], Verena Scheper[1], Minoo Lenarz[2], Thomas Lenarz[1], Andrej Kral[1] and Hubert H. Lim[3]

[1] *Department of Otorhinolaryngology, Hannover Medical University, Hannover, Germany*
[2] *Department of Otorhinolaryngology, Berlin Medical University – Charite, Berlin, Germany*
[3] *Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN, USA*

Edited by:
Catherine Carr, University of Maryland, USA

Reviewed by:
Ken K. L. Yung, Hong Kong Baptist University, Hong Kong
Christine Portfors, Washington State University, USA

****Correspondence:***
Thilo Rode, Department of Otorhinolaryngology, Hannover Medical University, Karl-Wiechert-Allee 3, 30625 Hannover, Germany
e-mail: rode.thilo@mh-hannover.de

The auditory midbrain implant (AMI) consists of a single shank array (20 sites) for stimulation along the tonotopic axis of the central nucleus of the inferior colliculus (ICC) and has been safely implanted in deaf patients who cannot benefit from a cochlear implant (CI). The AMI improves lip-reading abilities and environmental awareness in the implanted patients. However, the AMI cannot achieve the high levels of speech perception possible with the CI. It appears the AMI can transmit sufficient spectral cues but with limited temporal cues required for speech understanding. Currently, the AMI uses a CI-based strategy, which was originally designed to stimulate each frequency region along the cochlea with amplitude-modulated pulse trains matching the envelope of the bandpass-filtered sound components. However, it is unclear if this type of stimulation with only a single site within each frequency lamina of the ICC can elicit sufficient temporal cues for speech perception. At least speech understanding in quiet is still possible with envelope cues as low as 50 Hz. Therefore, we investigated how ICC neurons follow the bandpass-filtered envelope structure of natural stimuli in ketamine-anesthetized guinea pigs. We identified a subset of ICC neurons that could closely follow the envelope structure (up to ∼100 Hz) of a diverse set of species-specific calls, which was revealed by using a peripheral ear model to estimate the true bandpass-filtered envelopes observed by the brain. Although previous studies have suggested a complex neural transformation from the auditory nerve to the ICC, our data suggest that the brain maintains a robust temporal code in a subset of ICC neurons matching the envelope structure of natural stimuli. Clinically, these findings suggest that a CI-based strategy may still be effective for the AMI if the appropriate neurons are entrained to the envelope of the acoustic stimulus and can transmit sufficient temporal cues to higher centers.

Keywords: auditory brainstem implant, cochlear implant, envelope, inferior colliculus, model, phase locking, speech, temporal code

INTRODUCTION

An auditory midbrain implant (AMI) designed for stimulation across the central nucleus of the inferior colliculus (ICC) was implanted in deaf patients who could not sufficiently benefit from a cochlear implant (CI; Lim et al., 2007, 2009). These patients have neurofibromatosis type 2, which is a genetic disease that leads to bilateral acoustic neuromas. Removal of the tumors usually leads to complete damage of the auditory nerves, and thus it is not possible to transmit acoustic information to the brain using a CI, which is designed to stimulate the auditory nerve. The AMI has been implanted in five patients. Encouragingly, the AMI has shown to be safe for over 6 years and has improved environmental awareness and lip-reading capabilities for the implanted patients. However, open-set speech understanding remains significantly lower for the AMI compared to the CI. Based on previous studies in animals and humans (Lenarz et al., 2006; Lim et al., 2009, 2013; Calixto et al., 2012; McKay et al., 2013), it appears that the AMI can transmit sufficient spectral cues when the electrode array is aligned along the tonotopic gradient of the ICC; however, the AMI cannot transmit sufficient temporal cues shown to be important for speech understanding (Shannon et al., 1995, 2004). The AMI currently uses a stimulation strategy designed for CIs in which the sound is processed through bandpass filters and the envelope of each filtered signal is used to amplitude-modulate an electrical pulse train presented to a frequency-aligned site. The question arises as to whether such a CI-based strategy is appropriate for the AMI.

For the CI, it is possible to position electrode sites in different locations along the cochlea to transmit sufficient spectral cues while presenting varying temporal patterns on each site to sufficiently transmit temporal cues for speech understanding (Shannon et al., 1995, 2004). In contrast, the ICC is a three-dimensional structure consisting of two-dimensional isofrequency laminae (Geniec and Morest, 1971; Malmierca et al., 1995; Oliver, 2005). Although sites can be positioned in each lamina to convey spectral cues, it is unclear if an amplitude-modulated pulse pattern delivered to only one site in each lamina would transmit sufficient temporal cues. Previous animal studies have shown complex and

varying coding properties, including differences in temporal following abilities, for neurons along an ICC lamina (Semple and Aitkin, 1979; Stiebler, 1986; Ehret, 1997; Langner et al., 2002; Hage and Ehret, 2003; Holmstrom et al., 2010). Neurons sensitive to a similar frequency (i.e., with a similar best frequency, BF) can have dramatically different response patterns for the same stimulus. Furthermore, studies have shown that neurons within the ICC are no longer able to follow the fast fluctuations of a sound stimulus, particularly those above a few hundred hertz (Rees and Moller, 1987; Langner and Schreiner, 1988; Krishna and Semple, 2000; Frisina, 2001). Auditory nerve fibers can follow stimuli up to several thousand hertz (Johnson, 1980; Joris et al., 2004). Therefore, it has been proposed that temporal features have been converted, at least for faster components, into a rate code and/or population spiking pattern across ICC neurons, reflecting the complex transformations that have occurred through several nuclei from the auditory nerve (Langner and Schreiner, 1988; Langner et al., 2002; Joris et al., 2004; Wang et al., 2008; Huetz et al., 2011). These findings suggest that a CI-based stimulation strategy would not be appropriate for the ICC. Rather, electrode sites may need to be positioned fully across each lamina of the ICC and activated in various spatial and temporal patterns to transmit sufficient temporal cues. Implanting multiple shanks across the ICC to span each lamina increases surgical risk and identifying optimal ways to activate the numerous sites along each lamina would be challenging.

On the other hand, acoustic and CI studies in humans have shown that limited temporal cues, even down to 50 Hz, are sufficient for speech understanding in quiet (Shannon et al., 1995; Nie et al., 2006; Xu and Pfingst, 2008), which would be the initial goal for the AMI. ICC neurons are still capable of following the temporal pattern of sound stimuli above 50 Hz and even up to a few hundred hertz, at least for artificial click trains and amplitude modulated stimuli (Rees and Moller, 1987; Krishna and Semple, 2000; Frisina, 2001; Langner et al., 2002; Joris et al., 2004; Zheng and Escabi, 2008). Therefore, a CI-based strategy may be effective for the AMI, assuming that activation of just a subset of ICC neurons with a single site within each lamina is sufficient for restoring speech understanding. Although the AMI uses a CI-based strategy in the current patients, it is not clear if the limited performance is related to the stimulation strategy or if the AMI sites are just not activating the appropriate neurons in the ICC that would be able to transmit the necessary temporal cues to higher centers.

Previous animal studies investigating the envelope following capabilities of ICC neurons to natural stimuli generally investigated the overall response patterns averaged across neurons from different frequency and isofrequency locations (Suta et al., 2003; Woolley et al., 2006). These studies suggested that ICC neurons, as a population response, could approximately follow the envelope pattern of different vocalizations. One of these studies by Suta et al. (2003) also presented spiking responses for a few individual ICC neurons. However, further studies are needed to directly assess how well ICC neurons within different frequency regions follow the envelope structure of the vocalizations, and thus if it is possible to use a CI-based strategy for ICC stimulation. One challenge in achieving this goal is how to directly compare the ICC spiking response with the envelope structure of the stimulus. Sound is modified through the outer and middle ear and converted into basilar membrane motion at different frequency locations along the cochlea, undergoing several non-linear transformations (Rhode, 1971; Meddis et al., 2001; Sumner et al., 2003; Irino and Patterson, 2006; Jepsen et al., 2008). Thus, it would seem appropriate to directly compare the ICC responses to the envelope of this basilar membrane motion corresponding to the correct BF locations and incorporating the complex transformations through the peripheral system (e.g., frequency tuning characteristics along the basilar membrane, non-linear amplitude compression of the cochlea, and middle ear filtering). In addition to these peripheral effects, it is also possible to incorporate the transformations through the hair cell synapses and auditory nerve fibers to estimate the neural pattern that is observed at the input of the brain (Zhang et al., 2001; Bruce et al., 2003; Sumner et al., 2003; Zilany et al., 2009).

Based on the considerations presented above, we performed a detailed assessment in ketamine-anesthetized guinea pigs of how well ICC neurons across different frequency regions can follow the envelope structure of a diverse range of spectral and temporal features as found in species-specific vocalizations. We recorded neural activity across a large number of sites within the ICC by using multi-site electrode arrays to more effectively sample the different types of possible responses. We also incorporated a peripheral ear model specifically designed for the guinea pig that provides the basilar membrane signal for each frequency channel (i.e., velocity at a specific basilar membrane location; Meddis et al., 2001; Sumner et al., 2003). The envelope of this signal can then be compared to the responses recorded in the corresponding frequency region of the ICC. The effects associated with the hair cell synapses and auditory nerve fibers were not incorporated into our peripheral modeling because we were initially interested in treating the transformation from the basilar membrane to the ICC as a single black box without incorporating any "neural" components. In this way, we could directly assess if ICC neurons can accurately follow the BF-matched envelope of the stimulus that is actually driving the neural impulses into the brain. Later, we will incorporate additional components to more accurately reveal how information is coded from the basilar membrane to the ICC, expanding on models from previous studies (Hewitt and Meddis, 1994; Nelson and Carney, 2004; Guerin et al., 2006; Jepsen et al., 2008; Dugue et al., 2010). By incorporating a simple model of the peripheral ear, we were able to observe a surprisingly close resemblance between the BF-matched envelope of the stimulus and the spiking response of ICC neurons. There were many neurons that did not exhibit responses resembling the temporal structure of the stimulus. However, the existence of at least a subset of neurons spanning different frequency regions across the ICC that could closely follow the envelope of vocalizations suggests that CI-based stimulation strategies may still be successful for the AMI.

MATERIALS AND METHODS

Basic surgical procedures and methods for neural recording and acoustic stimulation were similar to those presented in previous work (Lim and Anderson, 2006; Calixto et al., 2012) and performed in ketamine-anesthetized guinea pigs. For this study, we

acoustically presented species-specific vocalizations to the animals and recorded the corresponding spike activity along the tonotopic axis of the ICC using a multi-site electrode array. We then plotted the temporal pattern of spiking activity as post-stimulus time histograms (PSTHs) and compared them with the different envelopes of the presented vocalizations. Each PSTH corresponds to a recording site within a given BF region. Each envelope of a vocalization corresponds to the motion (i.e., velocity in our study) of the corresponding BF location along the basilar membrane that is outputted from a dual resonance non-linear (DRNL) model, which was developed for the guinea pig peripheral auditory system by Meddis et al. (2001) and Sumner et al. (2003). The DRNL model estimates the transformation of the sound stimulus through the cochlea after it is passes through a set of middle ear filters, outputting a signal for each basilar membrane location. We define the DRNL output as the envelope of each of these BF-matched signals.

VOCALIZATION RECORDINGS

We recorded species-specific vocalizations from adult guinea pigs (Dunkin Hartley; Harlan Laboratories, Venray, Netherlands), which were the same breed of animals later used for the neurophysiological recording experiments. The different vocalizations were recorded using a calibrated omnidirectional condenser microphone (model MK202E, Microtech Gefell, Germany), amplified through a TDT MA3 microphone amplifier (Tucker-Davis Technology, Alachua, FL, USA), digitized with a TDT RX6 system (97.5 kHz sampling frequency, 24-bit resolution), and processed using custom Matlab software (MathWorks, Natick, MA, USA). To avoid background noise, the recordings were performed in a sound attenuating recording box (Human Tec, Leopoldshöhe/Greste, Germany). During processing of the recorded vocalizations, we only selected those that exhibited a high signal-to-noise ratio and with minimal interference from movement (determined through video recordings) or vocalizations from other animals. We also ensured that a 50 ms quiet period existed before and after each vocalization for the sound stimuli used in the experiments. A 25 ms cosine ramp was used at the onset and offset of each of these stimuli (i.e., during those 50 ms portions) to avoid any transient steps that could affect the neural responses recorded in the ICC. In addition, the signals were bandpass filtered between 500 Hz and 40 kHz, which corresponded to the effective range of our TDT CF1 speakers. Although there have been up to 11 types of calls documented for the guinea pig (Harper, 1976; Suta et al., 2003; Philibert et al., 2005; Grimsley et al., 2012), we recorded ICC response patterns to only three vocalizations in this study due to time limitations during each experiment (**Figure 1**). We selected these three vocalizations because they exhibit a wide range of spectral and temporal features characteristic of the different types of natural sounds generated by guinea pigs. The first one consists of a harmonic upward sweep with varying frequency components over time, which we call a temporally and spectrally varying stimulus (TSV). The second one contains two broadband components surrounding a shorter harmonic portion in the middle, which we call a broadband with harmonic components stimulus (BH). The third one consists of several short bursts that each cover a broad frequency range, which we call a broadband and transient stimulus (BT).

SOUND CALIBRATION

Pure tones, broadband noise, and vocalizations were presented through TDT CF1 speakers coupled to the left ear through a hollow ear bar. For calibration, the same setup was used with a 0.25-inch condenser microphone (ACO Pacific, Belmont, CA, USA) where the tip of the ear bar was inserted into a short plastic tube with the microphone inserted into the other end. The tube matched the dimensions of a typical guinea pig ear canal. Pure tones and broadband noise stimuli were calibrated by correcting for the magnitude response obtained from playing pure tones through the speaker–ear system. However, to calibrate the vocalizations, the phase response also had to be corrected. For this purpose, we applied a normalized least mean squares (NLMS) adaptive filter to estimate the complete inverse transfer function of the speaker–ear system (Moonen and Proudler, 1998). We calibrated each vocalization for different levels ranging from 30 to 70 dB SPL (10 dB steps) that corresponded to the peak value from a 20 ms sliding root mean square window of each signal.

ANESTHESIA AND SURGERY

All experiments were performed on male or female albino guinea pigs (458–749 g; Dunkin Hartley; Harlan Laboratories, Venray, NL). We initially anesthetized the animals with an intramuscular injection of ketamine (40 mg/kg) and xylazine (10 mg/kg) with periodic supplements to maintain a non-reflexive state. Atropine sulfate (0.05 mg/kg) was injected subcutaneously throughout the experiment to reduce bronchial secretion. A warm water heating blanket controlled by a rectal temperature probe was used to keep the body temperature at 38 ± 0.5°C. All experiments were carried out in accordance with the German law for animal protection and were approved by the regional government (Landesamtes für Verbraucherschutz und Lebensmittelsicherheit registration number 33.9-42502-04-09/1666).

Before surgery, we recorded auditory brainstem responses (ABRs) using 10 ms pure tones as acoustic stimuli (1, 4, 8, 16, and 32 kHz; 1 ms rise–fall ramp times) and low impedance subcutaneous electrodes positioned at the left mastoid (signal), vertex (reference), and frontal skull region (ground). We excluded any animals with hearing thresholds greater than 30 dB SPL at 8 kHz or abnormally shaped audiograms (typical examples shown in Gourevitch et al., 2009). After fixing the animal into a stereotaxic frame (David Kopf Instruments, Tujunga, CA, USA), we exposed the right occipital lobe. To position a multi-site electrode array (NeuroNexus Technologies, Ann Arbor, MI, USA) into the ICC, we partially aspirated the occipital cortex and exposed the right inferior colliculus surface (Bledsoe et al., 2003; Snyder et al., 2004). The array consisted of two shanks (500 μm separation) each with 16 linearly spaced sites (100 μm spacing, $\sim$400 μm^2 area). The array was inserted at a 45° angle to the sagittal plane with each shank aligned along the tonotopic axis of the ICC (Malmierca et al., 1995; Snyder et al., 2004; Lim and Anderson, 2006). The ground wire was placed in the neck muscles. After placement of the array, the brain was covered with agarose gel. During the experiment, the array was repositioned two to three times.

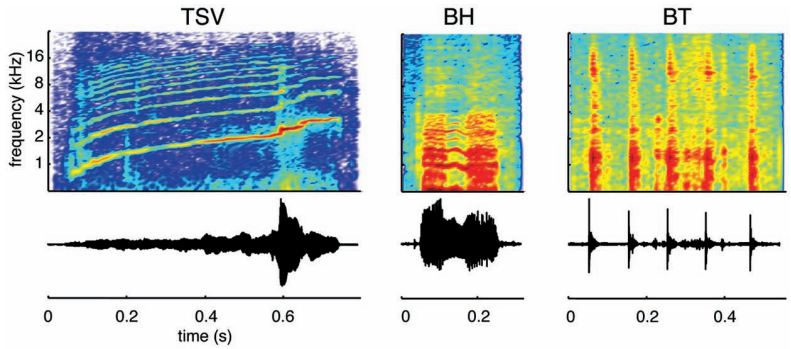

FIGURE 1 | Vocalization waveforms and spectrograms. Three different types of vocalizations were used as acoustic stimuli: a temporally and spectrally varying stimulus (TSV), a broadband with harmonic components stimulus (BH), and a broadband and transient stimulus (BT). All three cover a wide spectral range but differ in their spectral and temporal characteristics. The vocalizations presented to the animals were bandpass filtered (500 Hz to 40 kHz) and calibrated with respect to the speaker–ear interface.

All experiments used for this paper had a total duration between 12 and 16 h. As described above, ABRs were obtained at the beginning of each experiment, which required approximately 45 min. The surgery required about 2–3 h. The neural recordings from the ICC had a duration between 8 and 12 h. All animals were euthanized after each experiment by decapitation under deep anesthesia.

STIMULATION AND RECORDING SETUP

All experiments were performed in an acoustically and electrically shielded chamber and controlled by a computer interfaced with TDT System 3 hardware using custom software written in Matlab. All stimuli were calibrated as described above. To aid in positioning the electrode array, we presented various levels of pure tones and broadband noise that were 50 ms in duration with 5 ms and 0.5 ms rise–fall ramp times, respectively, to elicit acoustic-driven activity in the contralateral ICC. Once the array was correctly positioned within the ICC, each vocalization was presented at different levels ranging from 30 to 70 dB SPL in 10 dB steps and randomized across all stimuli (20 trials for each stimulus). Inter-stimulus intervals were at least 1.5 s long. We also recorded 20 trials of spontaneous neural activity randomly interleaved with the stimulus trials. All neural signals were passed through analog DC-blocking and anti-aliasing filters from 1.6 Hz to 7.5 kHz. The sampling frequency used for acoustic stimulation was 195 kHz and for neural recording was 24 kHz. The sampling rate of the vocalizations was doubled from 97.5 kHz to match the rate of 195 kHz used for the experiments.

PLACEMENT OF ELECTRODE ARRAY IN ICC

Post-stimulus time histograms and frequency response maps (FRMs) were plotted online to confirm that the electrode arrays were correctly positioned along the tonotopic axis of the ICC. Details on these analysis methods and example plots for similar types of electrode arrays are presented in previous publications (Lim and Anderson, 2006; Neuheiser et al., 2010). Briefly, for the FRMs we presented pure tones with a length of 50 ms with 5 ms rise–fall cosine ramps, levels ranging from 0 to 70 dB SPL in 10 dB steps, and frequencies ranging from 500 Hz to 50 kHz with six steps per octave. All stimuli were randomly interleaved. Neural recordings were obtained 20 ms before the onset of each stimulus and for a duration of 200 ms. We bandpass filtered the neural signals (300–3000 Hz) and detected spikes on each site that exceeded a manually adjusted threshold above the background activity. We binned the spikes into PSTHs (1 ms bins). The number of trials for broadband stimulation varied whereas four trials were presented for each pure tone and level stimulus for the FRMs. To create a FRM for each site, we calculated the driven spike rate (total minus spontaneous spike rate) within a set PSTH window relative to the stimulus onset (5–65 ms) and plotted that value for each frequency-level combination. Sustained PSTHs in response to broadband stimuli as well as a systematic shift in frequency tuning from low to high frequencies for superficial to deeper sites confirmed that our array was positioned within the ICC.

OFFLINE DATA ANALYSIS

For the analysis, data from 10 animals were used. The figures in this paper include either single cases or combined data from all 10 animals. We compared the multi-unit spiking activity recorded from different frequency regions of the ICC with the corresponding spectral and temporal pattern of the presented vocalizations. In order to obtain data from many different locations across ICC, we used multi-site arrays that generally record multi-unit activity. It was not possible to reliably isolate single-unit spikes from this multi-unit activity. However, since we are interested in assessing temporal patterns that would guide ICC stimulation, which generally activates clusters of neurons surrounding each electrode site, analysis of multi-unit responses was sufficient for this study. These multi-unit spikes were automatically detected at a threshold of 3.5 times the standard deviation of the background noise recorded on a given site after the signals were bandpass filtered between 300 and 3000 Hz. Only negative peaks were detected as spikes. The total spikes across trials minus the mean spike count across spontaneous trials were calculated for 1-ms bins to obtain a stimulus-driven PSTH for each site and stimulus (i.e., a vocalization at a certain level). We reprocessed the FRM spike data offline in a similar way and calculated the BF by taking the centroid frequency value

at 10 dB above the visually determined threshold level from the FRMs, as performed in our previous studies (Lim and Anderson, 2006; Neuheiser et al., 2010). We used this BF measure instead of characteristic frequency (i.e., frequency corresponding to the maximum activity at threshold) because it was less susceptible to noise and more consistent with what we visually estimated from the FRMs.

The PSTH for each site with a specific BF and for a given vocalization was then compared with its BF-matched DRNL output. The DRNL output corresponds to the envelope of the basilar membrane velocity at a given BF location. The DRNL model of the peripheral auditory system was developed by Meddis et al. (2001) and modified for the guinea pig by Sumner et al. (2003; **Figure 2**). The DRNL model uses stapes velocity as the input. To obtain the stapes velocity, we first passed the signal through a middle ear filter, which consisted of a cascade of two Butterworth bandpass filters. One was a second order filter with a bandpass range from 4 to 25 kHz, while the other was a third order filter with a bandpass range from 700 Hz to 30 kHz. Both had unity gain. This middle ear filtering was necessary to reproduce the appropriate activation thresholds observed at the basilar membrane and auditory nerve fibers in the guinea pig (Evans, 1972; Sumner et al., 2003). The output from the middle ear filter was then scaled by a factor of 1.4×10^{-10} to convert sound pressure into stapes velocity (in m/s). Further details on this middle ear filtering are provided in Sumner et al. (2003). The DRNL model has both a non-linear and a linear path in parallel. The linear path consists of a linear gain, a cascade of three first order gammatone filters, and a cascade of four second order Butterworth lowpass filters. The non-linear path consists of a cascade of three first order gammatone filters,

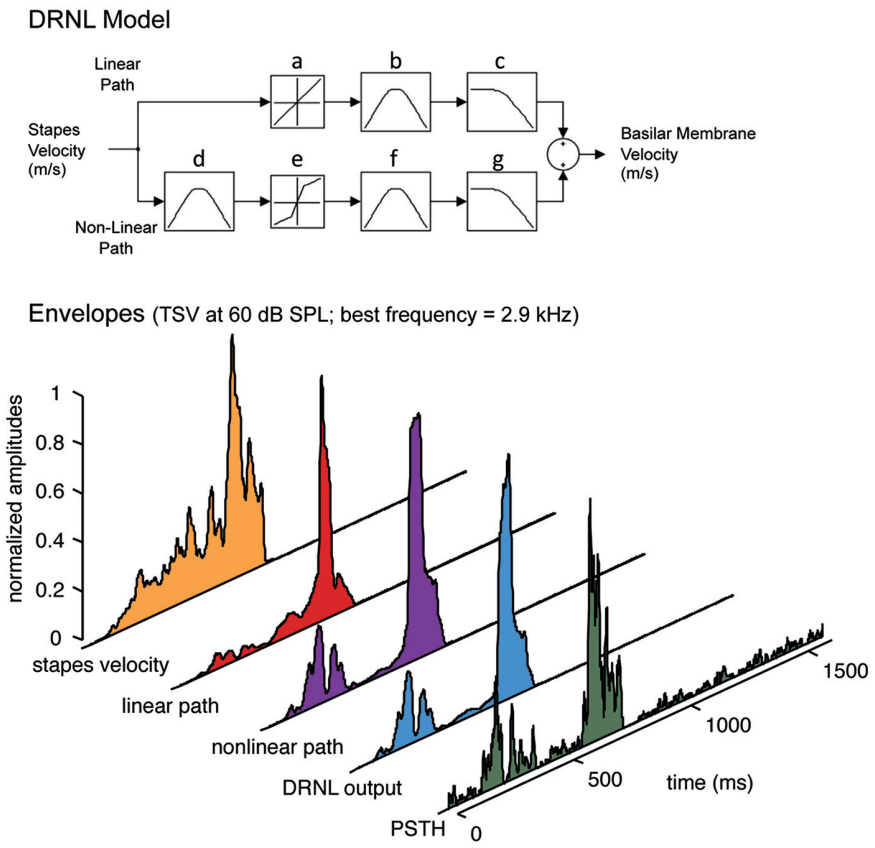

FIGURE 2 | Peripheral ear modeling and data processing. *Top*: The dual resonance non-linear (DRNL) model is fit with guinea pig parameters. The model simulates the peripheral processing from the stapes input to the output of a specified best frequency location along the basilar membrane. The model consists of a linear and a non-linear path, each with a set of gammatone bandpass filters (b, d, f) and Butterworth lowpass filters (c, g) that vary depending on the specified best frequency. In addition, the non-linear path of the model contains a broken stick compression function (e; also varying with best frequency and in contrast to the linear function in a) to estimate the non-linear effects of the cochlea. *Bottom*: The signals are extracted for each part of the DRNL model and compared with the neural response for a single ICC site. Before applying the DRNL model, the signal is first inputted through a middle ear filter (two Butterworth bandpass filters from 4 to 25 kHz and from 700 Hz to 30 kHz) and scaled from sound pressure to *stapes velocity*. The output of the DRNL model is obtained from summing the *linear path* and *non-linear path* components. The envelope of the DRNL model (i.e., what we define as the *DRNL output*) is then extracted using the Hilbert transform method and smoothed by a 10 ms sliding window average to obtain the slower fluctuations of the signal (see Materials and Methods for further details). The envelopes for the other signal components were extracted in the same way and displayed in the figure for visualization. The post-stimulus time histogram (PSTH), which was also smoothed with a 10 ms sliding window average, shows the neural activity recorded in a 2.9 kHz frequency lamina in response to TSV shown in **Figure 1**. The best frequency of 2.9 kHz was used for the DRNL model to obtain the different envelope signals shown in the figure, except for stapes velocity, which is independent of best frequency.

a non-linear broken-stick transfer function to account for compression effects, another cascade of three first order gammatone filters, and a cascade of four first order Butterworth lowpass filter. The outputs of the linear and non-linear paths are then summed together.

Further details on the different parameters and values used for the DRNL model are presented in Sumner et al. (2003) and a freely available Matlab program was obtained from Drs Ray Meddis and Christian Sumner to perform the DRNL simulation in our study. In brief, several parameters were fixed while other parameters changed as a function of BF. To estimate the basilar membrane velocity at a given location, the Matlab program required only the corresponding frequency (i.e., referred to BF in our study) and the vocalization waveform scaled to the desired sound pressure level. The program then calculated and implemented the corresponding bandwidths, cut-off frequencies, gains, and scaling factors required for each component of the DRNL model. This DRNL model has been able to reproduce various tuning properties and non-linear input–output functions at different locations along the basilar membrane, variations in phase response with frequency and level, impulse responses on the basilar membrane, local distortion products, and other peripheral phenomenon that have been measured across several species (Lopez-Poveda and Meddis, 2001; Meddis et al., 2001; Sumner et al., 2003). Combining the DRNL model with an inner hair cell model was also able to reproduce the BF dependence of frequency–threshold tuning curves, rate-intensity functions, iso-intensity functions, compression effects, and basic temporal firing patterns to speech stimuli as measured in the auditory nerve fibers in guinea pig (Sumner et al., 2003; Holmes et al., 2004). As explained in Section "Introduction" for our study, we used the DRNL model without the inner hair cell model because we were initially interested in treating the transformation from the basilar membrane to the ICC as a single black box without incorporating any "neural" components.

Once the output signal from the DRNL model was calculated, we extracted its envelope using the Hilbert transform method, which has shown to successfully extract perceptually relevant envelope features for speech understanding (Smith et al., 2002; Nie et al., 2006). The envelope was downsampled to a resolution of 1 ms to match that of the PSTH. We then further smoothed both the envelope (to obtain what we define as the "DRNL output") and PSTH using a 10 ms sliding window average, which keeps envelope fluctuations up to ∼100 Hz. All PSTH responses and DRNL outputs were normalized by their maximum value (i.e., normalized to 1) for analysis and improved visualization. **Figure 2** presents the normalized envelopes of the different signal components of the model all smoothed with the 10 ms window average to enable direct comparison with the corresponding PSTH. We selected a 10 ms window because we were interested in envelope fluctuations relevant for speech understanding (at least 50 Hz; Shannon et al., 1995) yet within the range of temporal following capabilities shown for ICC neurons, which are mostly below or around 100 Hz across several species, including guinea pig (Rees and Moller, 1987; Rees and Palmer, 1989; Snyder et al., 2000; Joris et al., 2004). Based on visual inspection of the data, we also observed more noticeable differences between the PSTHs and DRNL outputs once we began to decrease the window below 10 ms.

To directly assess similarity, we calculated the peak of the cross-correlation function between a PSTH response and its corresponding DRNL output. If we observed any peak values at a time lag less than 5 ms or above 20 ms, which is typically outside the range for acoustic activation of the ICC (Syka et al., 2000), we visually checked and manually corrected for any artificial correlation values. Additionally, we excluded any cases in which there was low driven activity below a set threshold that prevented reliable comparison with the DRNL output. This set threshold corresponded to four times the standard deviation of the PSTH activity from the spontaneous trials for each recording site. Only PSTHs that consisted of more than 20 bins above that set threshold during the stimulation period were considered valid cases, which matched what we determined through visual inspection of the data. All the PSTHs analyzed in this study spanned a BF range of 500 Hz to 20 kHz. Unless stated in the text or figures, all analyses and plots correspond to a sound level of 60 dB SPL.

RESULTS

Data were collected from 10 animals. The total number of recording sites in ICC varied for each vocalization and will be specified when presenting the different results.

ICC RESPONSES TO VOCALIZATIONS

A wide range of responses was recorded in the ICC in response to the three different vocalizations. **Figure 3A** presents one example for each vocalization corresponding to one of the highest correlation values. The DRNL output (shaded region) for BT corresponding to a BF of 12.3 kHz is quite similar to the corresponding PSTH response (black line). The PSTH slightly lags behind the DRNL output as expected for the acoustic activation to reach the ICC. The PSTH response was also quite similar to the DRNL output for BH. For TSV, the PSTH response followed almost every peak in the DRNL output, being partially misaligned in magnitude for a few peaks, which is still remarkable considering the large number of peaks and complex pattern of the stimulus envelope compared to those for BT and BH. On the other extreme, we observed PSTH responses that were quite different from the DRNL output in which the temporal patterns and peaks were completely misaligned. **Figure 3B** presents one example for each vocalization with the lowest correlation value across our entire data set. These results suggest that some ICC neurons can closely follow the envelope structure of natural stimuli, while other neurons exhibit different spiking patterns not locked to the stimulus envelope.

The cross-correlation provides a measure of the similarity between two curves. A value of one indicates a perfect match between two curves. To provide a qualitative interpretation of the different correlation values observed in our study, we present eight typical examples for each vocalization with the maximum (**Figure 4**) and median (**Figure 5**) correlation values we observed for different frequency regions across our data set. As shown for the maximum correlation cases (for those roughly above 0.9), there are ICC neurons that can generally follow the envelope structure of the basilar membrane in response to different vocalizations and across a wide range of BF regions. The greatest similarity was observed for the BT examples. Some cases for BH and TSV

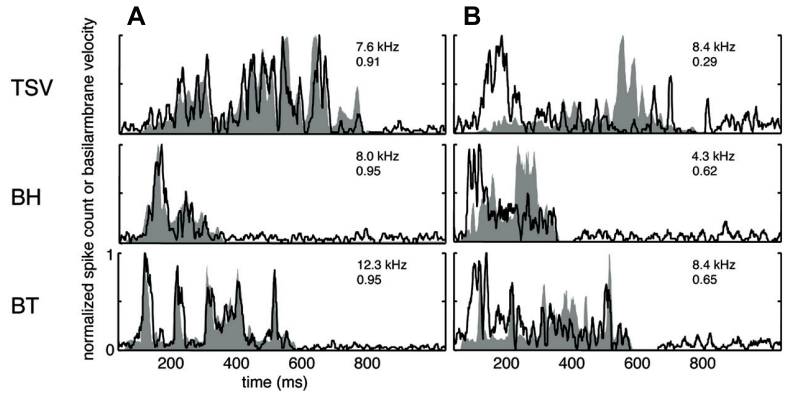

FIGURE 3 | Correlation examples. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for each recording site are labeled on the top-right of each plot. **(A)** High correlation examples are shown for the three vocalizations presented in **Figure 1**. **(B)** The lowest correlation examples across our data set are shown for the three vocalizations.

exhibited greater variability between the DRNL output and the PSTH response. However, all of the displayed PSTH responses generally followed the temporal timing of most of the different peaks of the DRNL output, with deviations mainly related to the actual amplitudes of those peaks. The DRNL model provides an estimate of the basilar membrane motion and was fitted with a single set of parameters to represent all BF regions. It is possible that these amplitude discrepancies could be in part due to sub-optimal modeling of the amplitude non-linearities at the basilar membrane and would require better fitting of the parameters for each BF region separately. Even if the examples shown in **Figure 4** are an underestimation of the actual similarity between the DRNL outputs and PSTH responses, they still demonstrate that there are at least some ICC neurons that can closely or approximately follow the envelope structure of natural stimuli.

The median examples across BF regions are presented in **Figure 5**. The median correlation value was 0.75 ($n = 239$), 0.87 ($n = 203$), and 0.86 ($n = 277$) for TSV, BH, and BT, respectively. Median PSTH examples for BH and BT can approximately follow the DRNL outputs though there are greater differences between the peak amplitudes compared to those in the maximum correlation group. The differences in temporal timing and amplitude of the peaks between the PSTHs and DRNL outputs are greater for TSV than for the other two calls.

Based on all the different ICC examples from our data set, including the typical ones shown in **Figures 4** and **5**, correlation values above 0.85 correspond to ICC responses that approximately or closely follow the DRNL output. In particular, 15% (36 of 239), 60% (122 of 203), and 58% (160 of 277) of the sites for TSV, BH, and BT, respectively, satisfy this criterion. Out of the 36 TSV cases with high correlation values ($\geq$0.85), 64% had high correlation values also for BH and BT, while 78% had high correlation values for at least one of the other calls. Additionally, 75% (92 of 122) of the BH cases with high correlation values also had high correlation values for BT. Therefore, the same ICC neurons can follow the envelope structure for two or all three of the vocalizations used in our study demonstrating that there is a subset of ICC neurons designed to temporally code for a diverse range of spectral and temporal patterns relevant for natural sound processing. These neurons generally follow the envelope structure for BH and BT better than for TSV, which may relate to the more complex and time-varying spectral pattern of TSV.

SPECTRAL AND TEMPORAL FEATURES

With the two-shank array, it was possible to simultaneously record neural activity from two sites within a similar isofrequency lamina of the ICC in addition to multiple sites across different frequency regions. **Figures 6A, 7A,** and **8A** show data recorded from four different frequency laminae from the same sites and animal in response to the three different vocalizations. The DRNL output is approximately similar across the two sites in each lamina (i.e., slightly different due to slightly varying BF values). However, the PSTH responses could exhibit more drastic differences between the two sites even though they had approximately similar BF values. It appears that some peaks in the DRNL output are better represented by one site versus the other site, and vice versa. For example in **Figure 6A** for the 12.6–12.7 k Hz lamina, the site in the right column exhibited a PSTH response that more closely followed the peaks at ~0.6–0.7 s compared to the site in the left column. In addition, as mentioned above, the PSTH response more closely matched the DRNL output for BT and somewhat for BH compared to TSV. For example, this comparison can be observed for the same 6.5 kHz site in **Figures 6A** (TSV), **7A** (BH), and **8A** (BT).

For this study, we did not systematically investigate the effects of different locations along a given ICC lamina on temporal following properties. However, we pooled data across all our sites within and across different frequency regions to assess if the population PSTH could resemble the spectrotemporal patterns observed in the spectrogram for each vocalization. **Figure 6B** displays the spectrogram for TSV. In **Figure 6C**, the DRNL outputs across different frequency regions ranging from 500 Hz to 20 kHz were concatenated to create a plot that resembles a spectrogram. By comparing **Figures 6B,C**, it can be seen that the middle ear filtering and DRNL model attenuates the lower frequency components of the stimulus (i.e., relatively enhances the higher frequency components) while slightly smearing the stimulus in the frequency dimension

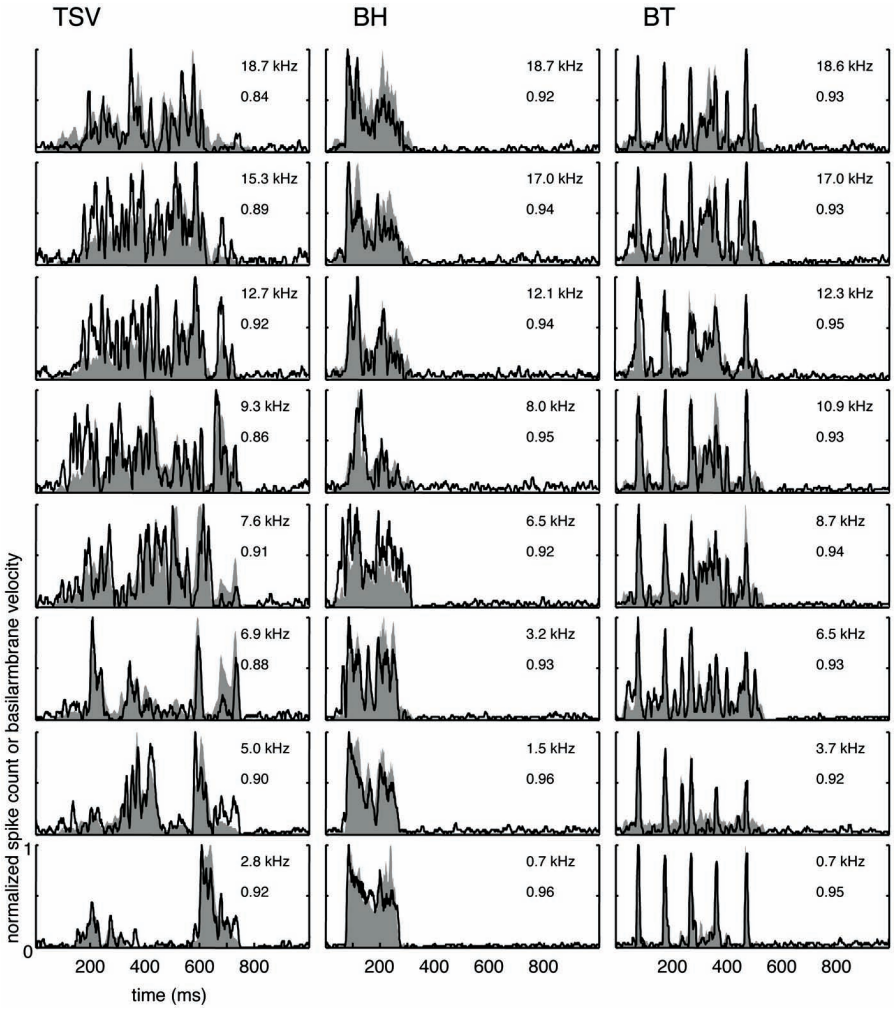

FIGURE 4 | Maximum correlations. Examples with the highest correlation values across our data set for different frequency regions are shown for the three vocalizations. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for each recording site are labeled on the top-right of each plot.

due to the overlapping frequency bands along the basilar membrane. The stimulus has also been smoothed in the time domain by the Hilbert transform method and a 10 ms sliding window average. In **Figure 6D**, we concatenated all the PSTHs recorded across the ICC in response to TSV at 60 dB SPL and obtained the population PSTH. The population PSTH more closely matches the DRNL spectrogram (**Figure 6C**) than the actual stimulus spectrogram (**Figure 6B**), especially for the weaker activity within the lower frequency regions. Similarly, general spectrotemporal features in the DRNL spectrogram can be observed in the population PSTH for both BH (**Figure 7**) and BT (**Figure 8**). Note that we did not have many sites for BFs below 2 kHz and thus cannot interpret how well ICC neurons follow the sound envelope for those lower frequencies. Interestingly, a few low BF sites had strong activity even though there was weak sound energy in the DRNL output. Strong activity in low frequency regions of the ICC associated with low stimulus energy has also been observed in a previous study (Suta et al., 2003), and may reflect the non-linear transformations that have occurred from the periphery up to ICC.

AMPLITUDE NON-LINEARITIES IN ICC ACTIVITY

The DRNL model incorporates amplitude compression through the non-linear path in which smaller amplitude components can increase to a greater extent than larger amplitude components, as has been observed in basilar membrane and auditory nerve activity in animals (Yates et al., 1990; Cooper and Yates, 1994; Meddis et al., 2001; Sumner et al., 2003). To assess the contribution of this compression component, we compared the DRNL output with the linear path output (the red component shown for the envelopes in **Figure 2**). The first column of **Figure 9** presents the DRNL output for the different vocalizations at several levels while the second column displays just the linear path output. Note that the linear path output is identical across levels for a given vocalization except for a scaling factor, which can be seen as the *y*-axis range changes for each level. All amplitude values change

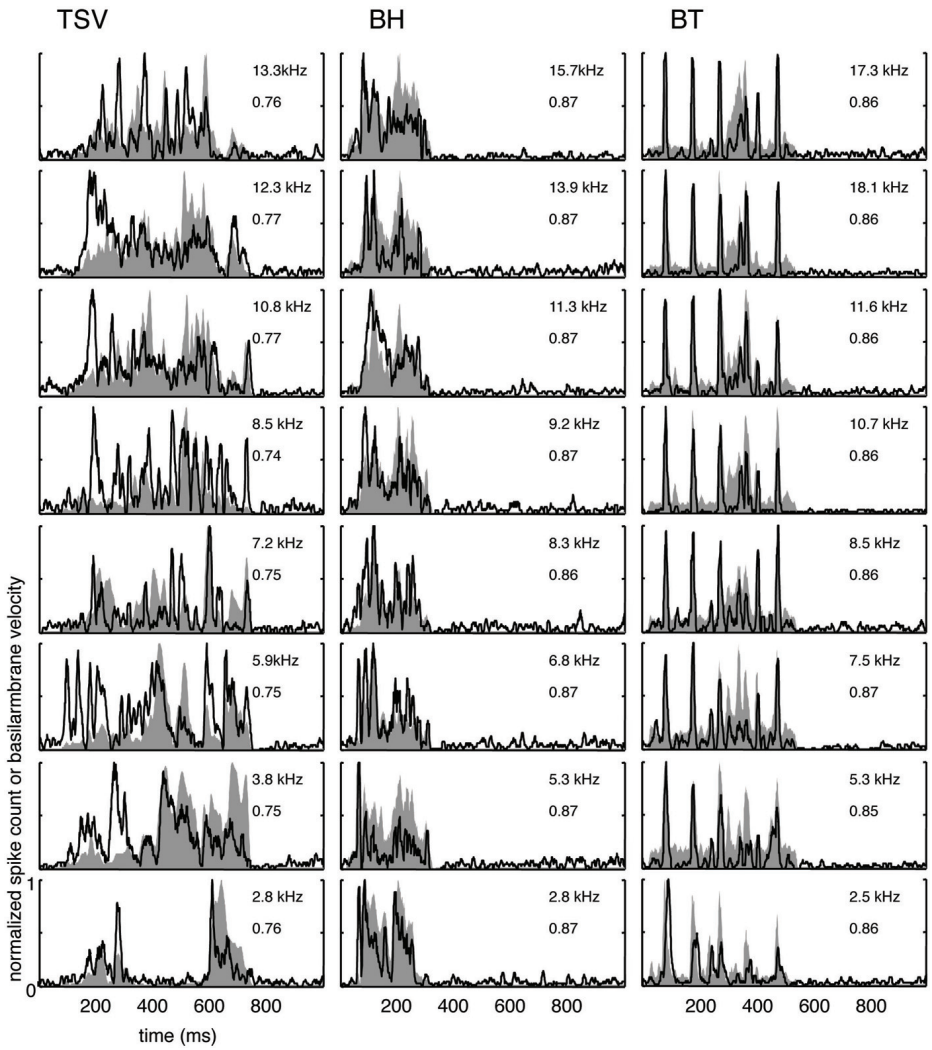

FIGURE 5 | Median correlations. Examples with correlation values around the median value across our data set are shown for different frequency regions for the three vocalizations. The median correlation value is 0.75 for TSV, 0.87 for BH, and 0.86 for BT. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for each recording site are labeled on the top-right of each plot.

proportionally within the linear path. For TSV, as the stimulus level increases from 50 to 70 dB SPL, the smaller PSTH values of the linear path (e.g., those between ∼100 and 400 ms) increase to a greater extent than the larger PSTH values. The DNRL model compensates for some of these non-linear effects. However, the changes in amplitude with increasing level do not appear to be fully captured by the DNRL model. Different types of compressive effects are also observed for the BH and BT examples. For BH, the DRNL model produces a pattern that generally matches the PSTH response better than that of the linear path (e.g., the peaks around 200 ms), but not for all of the peaks. For BT, there is also a mixture of effects in which the DRNL model performs better for some peaks while the linear path performs better for other peaks. Overall, the DRNL model is not able to fully produce the envelope patterns observed for all of our recorded ICC sites across different levels. It is important to note, as shown in the next section, that there are still some ICC neurons that can closely follow the DRNL output for different levels. However, we present the complex patterns in **Figure 9** to show that there are neural features that are not being fully predicted by the DRNL model.

The DNRL model has provided good estimates for the compression effects observed for the basilar membrane and auditory nerve in guinea pig (Meddis et al., 2001; Sumner et al., 2003). We also observed high correlation values between the PSTHs and DRNL outputs for multiple ICC sites (**Figures 3A** and **4**) indicating that some ICC neurons are following the motion at the basilar membrane predicted by the DNRL model. Therefore, the complex changes in response amplitude with increasing level in ICC neurons that are not predicted by the DRNL model could be associated with actual neural transformations that have occurred from the periphery up to the ICC.

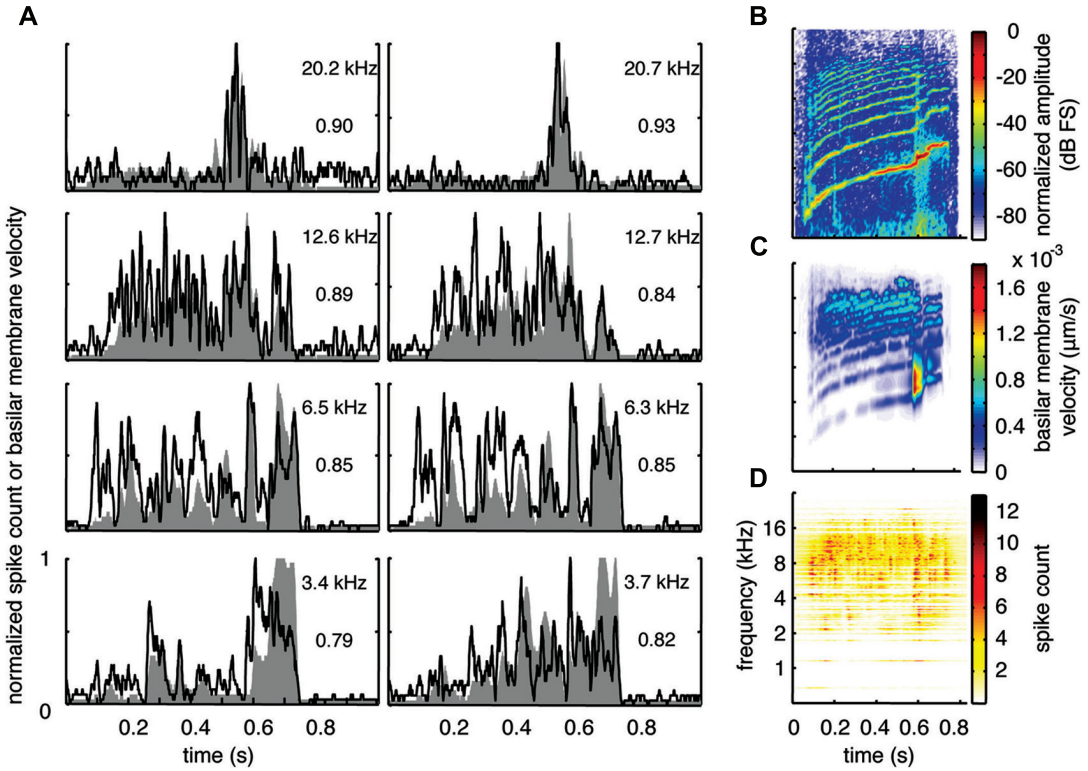

FIGURE 6 | Spectrotemporal representation for TSV. Neural activity was recorded across a wide range of best frequency regions from 10 animals and compared to the corresponding frequency-specific DRNL outputs of the stimulus. **(A)** DRNL outputs and PSTHs for four different frequency laminae (shown in rows) corresponding to two different shanks (shown in columns) from the same recording probe. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for each recording site are labeled on the top-right of each plot. **(B)** Spectrogram of the stimulus using the short-time Fourier transform. **(C)** Spectrogram made from the frequency-specific envelopes outputted from the DRNL model. Each row of the plot is an envelope outputted from one discrete DRNL frequency channel. Note that the different frequency channels of the DRNL model overlap in frequency ranges and thus smear the spectrogram plot in the frequency dimension. **(D)** Neural representation of the spectrotemporal features of activity recorded across all 239 sites. Each row is a PSTH recorded from each ICC site for a specific best frequency region. dB FS decibels relative to full scale in which 0 dB FS corresponds to the maximum value.

SUMMARY OF CORRELATION VALUES

In **Figure 10**, the correlation values for all our recording sites are plotted for each vocalization as a function of BF. The correlation values for BH and BT show consistently high values over the entire range of BFs and with values typically above 0.7. In contrast, there was a much wider range of correlation values for TSV, especially for lower BF sites. There appears to be a greater number of low BF neurons that are not capable of following the envelope for TSV even though many of those same neurons could reliably follow the envelopes for BH or BT.

We re-plotted the correlation values from **Figure 10** in a histogram format independent of BF to enable comparison with different stimulus levels with and without the non-linear path (**Figure 11**). For TSV, the DRNL output exhibited significantly higher correlation values compared to those of just the linear path at all three levels in which there was greater significance for higher levels. The BT also exhibited significantly higher correlation values for the DRNL output compared to those of the linear path but only at 40 and 50 dB SPL. In contrast to TSV, the significance values for BT actually decreased with increasing level such that the correlation values were significantly lower for the DRNL output versus the linear path output at 60 dB SPL. None of the levels for BH exhibited any significant differences in correlation values between the DRNL output and the linear path output. It is unclear as to why adding the non-linear component of the DRNL model improved the similarity between the DRNL outputs and the PSTH responses for TSV and BT but not for BH. TSV exhibits time-varying temporal and spectral patterns while BT exhibits more transient and broadband patterns. Considering that BH consists of a combination of patterns found in both TSV and BT, we also expected significantly higher correlation values for the DRNL output versus the linear path output for BH, at least for some levels.

In the previous section, we showed that ICC neurons could exhibit complex neural patterns across different levels that were not fully produced by the DRNL model. As shown in **Figure 11**, there are still ICC sites that exhibit high correlation values across different levels (i.e., for 40, 50, and 60 dB SPL). Due to an insufficient number of data points, we did not include analysis for 30 and 70 dB SPL in **Figure 11**. Interestingly, there was a moderate percentage of ICC sites that exhibited high correlation values ($\geq$0.85) across all three levels for BH (38.5%; 10 out of 26) and

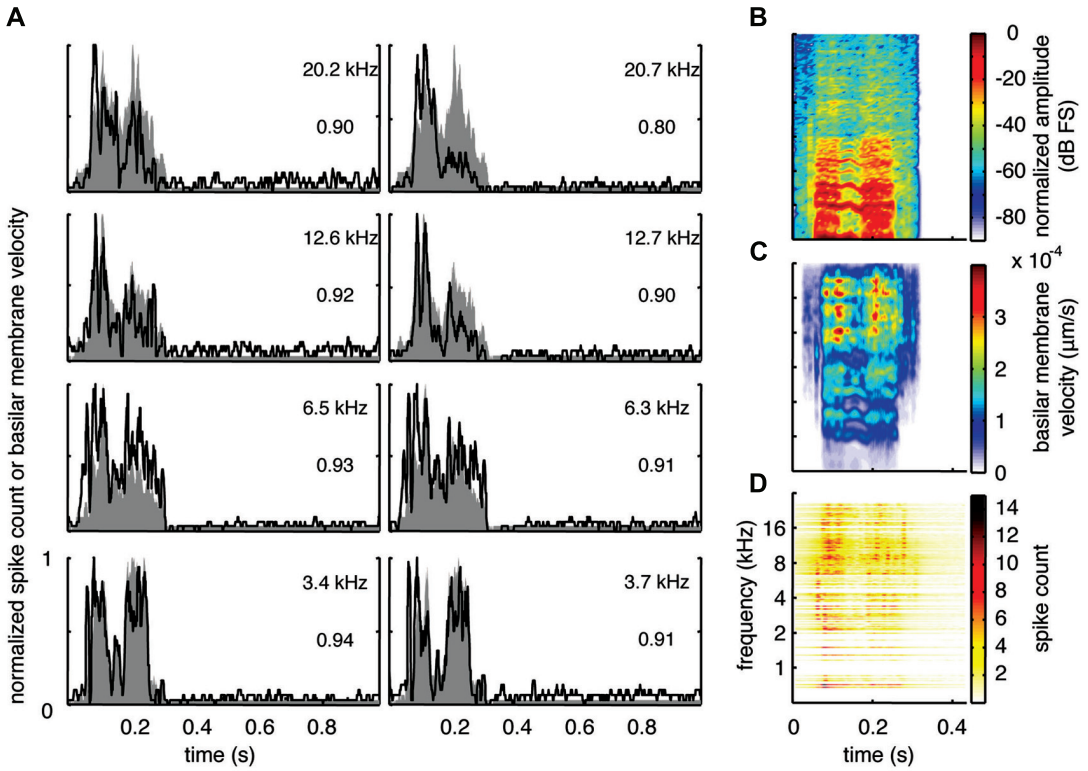

FIGURE 7 | Spectrotemporal representation for BH. Neural activity was recorded across a wide range of best frequency regions from 10 animals and compared to the corresponding frequency-specific DRNL outputs of the stimulus. **(A)** DRNL outputs and PSTHs for four different frequency laminae (shown in rows) corresponding to two different shanks (shown in columns) from the same recording probe. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for each recording site are labeled on the top-right of each plot. **(B)** Spectrogram of the stimulus using the short-time Fourier transform. **(C)** Spectrogram made from the frequency-specific envelopes outputted from the DRNL model. Each row of the plot is an envelope outputted from one discrete DRNL frequency channel. Note that the different frequency channels of the DRNL model overlap in frequency ranges and thus smear the spectrogram plot in the frequency dimension. **(D)** Neural representation of the spectrotemporal features of activity recorded across all 203 sites. Each row is a PSTH recorded from each ICC site for a specific best frequency region. dB FS decibels relative to full scale in which 0 dB FS corresponds to the maximum value.

BT (56.6%; 64 out of 113). For TSV, only 4.5% (4 out of 88) of the sites exhibited high correlation values across all three levels. It is important to note that each ICC site may not have data for all three levels, and thus the total number of sites for each vocalization listed in parentheses above is smaller than the total number (i.e., for *N*) listed in **Figure 11** for each level. Overall, these data suggest that there is at least a small subset of ICC neurons that are capable of coding for the envelope of the vocalizations across different levels as predicted by the DRNL model.

DISCUSSION

The three vocalizations used in this study were specifically selected to ensure that the neurons were activated with a broad spectrum of temporal and spectral patterns that the animals typically encounter on a daily basis. Based on our results, ICC neurons exhibit a wide range of neural patterns in response to these species-specific vocalizations. Previous studies have also shown that ICC neurons exhibit a wide range of neural patterns in response to artificial and natural stimuli (Rees et al., 1997; Krishna and Semple, 2000; Suta et al., 2003; Zheng and Escabi, 2008; Holmstrom et al., 2010) suggesting that different and complex transformations have occurred from the auditory nerve through the central auditory system. Cochlear neurons can reliably follow the temporal pattern of an acoustic stimulus, even up to several thousand hertz (Johnson, 1980; Joris et al., 2004). In contrast, ICC neurons generally follow the temporal patterns up to only a few hundred hertz (Rees and Moller, 1987; Krishna and Semple, 2000; Frisina, 2001; Langner et al., 2002). It has been proposed that the fast features of the stimulus have been converted into a complex rate code and/or population spiking pattern across neurons (Langner and Schreiner, 1988; Joris et al., 2004; Wang et al., 2008; Huetz et al., 2011). For the slower temporal components representing the envelope structure of the stimuli, there also appears to be a rate and/or place code in which different neurons are most sensitive to different amplitude modulation rates (Langner and Schreiner, 1988; Langner et al., 2002). Therefore, it is evident that the ICC is not simply replicating what occurs at the auditory nerve level, especially after passing through several synapses along the ascending auditory pathway. However, these transformations do not mean that there are no ICC neurons that can still reliably or approximately follow the envelope structure of natural stimuli. In this study, we present several cases in which ICC neurons closely followed the frequency-specific

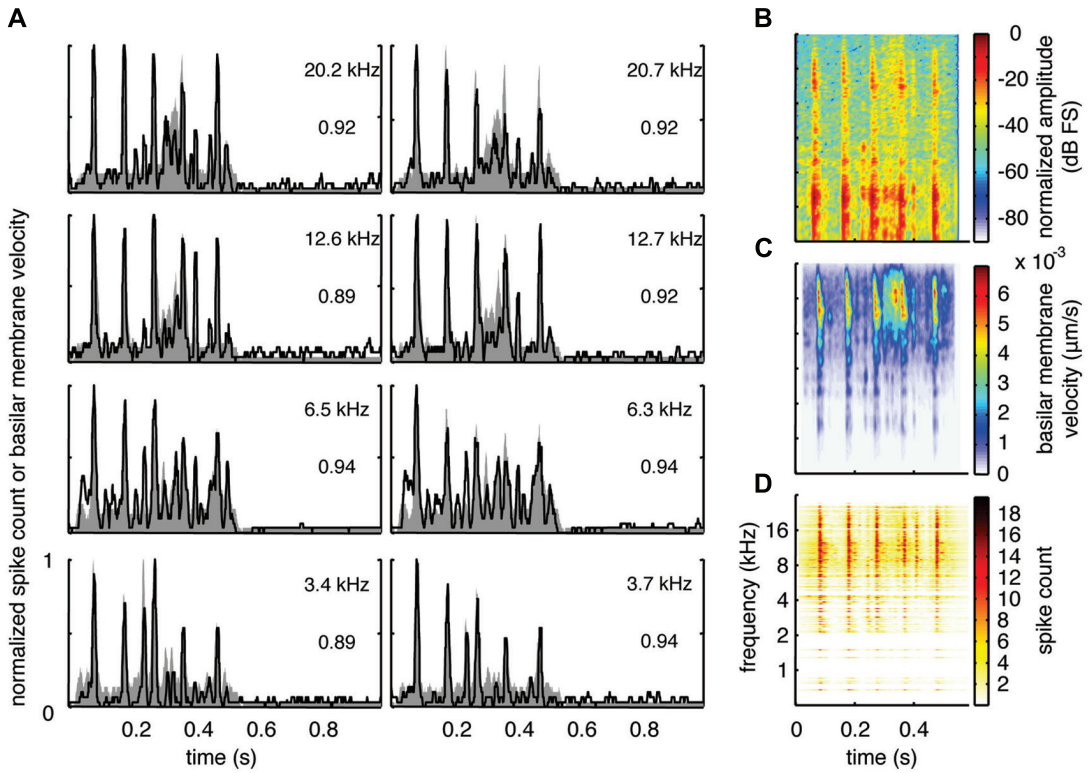

FIGURE 8 | Spectrotemporal representation for BT. Neural activity was recorded across a wide range of best frequency regions from 10 animals and compared to the corresponding frequency-specific DRNL outputs of the stimulus. **(A)** DRNL outputs and PSTHs for four different frequency laminae (shown in rows) corresponding to two different shanks (shown in columns) from the same recording probe. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for each recording site are labeled on the top-right of each plot. **(B)** Spectrogram of the stimulus using the short-time Fourier transform. **(C)** Spectrogram made from the frequency-specific envelopes outputted from the DRNL model. Each row of the plot is an envelope outputted from one discrete DRNL frequency channel. Note that the different frequency channels of the DRNL model overlap in frequency ranges and thus smear the spectrogram plot in the frequency dimension. **(D)** Neural representation of the spectrotemporal features of activity recorded across all 277 sites. Each row is a PSTH recorded from each ICC site for a specific best frequency region. dB FS decibels relative to full scale in which 0 dB FS corresponds to the maximum value.

envelope of different vocalizations, at least down to a smoothing window of 10 ms (up to ∼100 Hz). The following capabilities were greater for BT and BH compared to TSV. Interestingly, a large percentage of ICC neurons that followed the TSV stimulus with high correlation values were also able to closely follow the envelope structure of the other two calls. Based on these findings and previous studies, we propose that there are parallel pathways through the ICC and up to the auditory cortex in which some ICC neurons exhibit complex and transformed spiking patterns while other neurons still transmit a robust temporal representation matching the envelope of a diverse range of complex stimuli. The functional interpretation and supporting evidence for this view is further described below after discussing a few limitations of the study. We also discuss the implications of these findings for AMI implementation.

METHODOLOGICAL LIMITATIONS

There are a few limitations associated with the methods of our study that need to be discussed when interpreting the results. These limitations are associated with the use of a peripheral ear model to estimate the frequency-specific envelope at the basilar membrane, the effects of anesthesia, and the use of multi-unit activity versus single-unit responses.

To obtain the envelope of the basilar membrane at a specific BF location, we estimated the effects of the peripheral ear with a set of middle ear filters, a DRNL model, and an envelope extractor based on the Hilbert transform and a 10 ms smoothing function. The middle ear and DNRL components have been specifically designed for the guinea pig and have been able to estimate many features observed in physiological data: frequency tuning, nonlinear input–output functions, variations in phase response with frequency and level, and impulse responses at different BF locations along the basilar membrane; frequency–threshold tuning curves, rate-intensity functions, iso-intensity functions, compression effects for different BF auditory nerve responses; temporal spiking patterns of different BF auditory nerve fibers in response to single and double vowels; and other complex peripheral effects (e.g., two-tone suppression and local distortion products; Meddis et al., 2001; Sumner et al., 2003; Holmes et al., 2004). The Hilbert transform has also been shown to be a successful method for extracting out perceptually relevant envelope features for speech understanding (Smith et al., 2002; Nie et al., 2006). However,

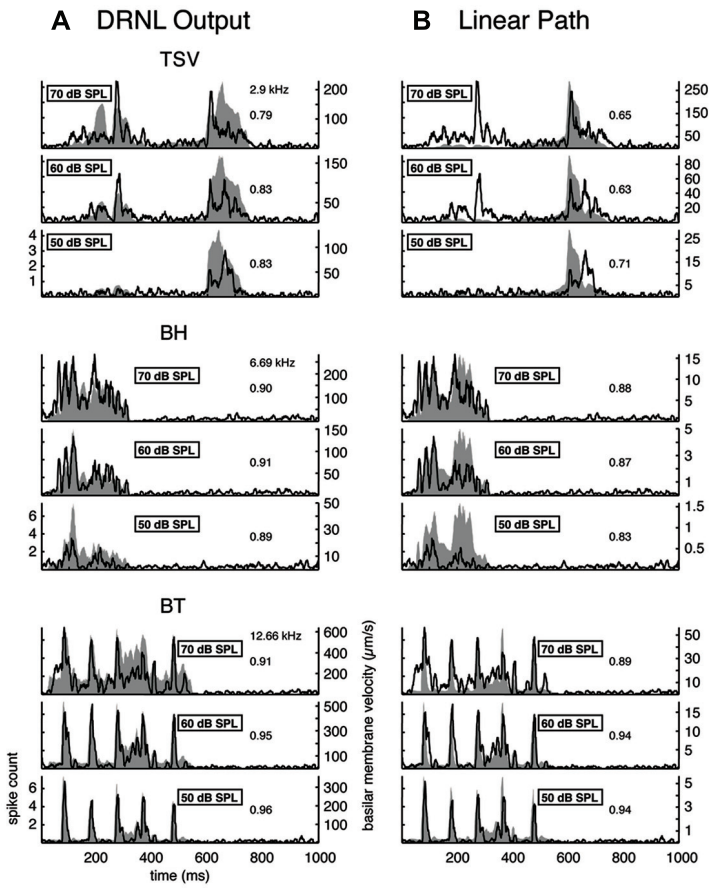

FIGURE 9 | Comparison of DRNL output versus the linear path output. The PSTH responses are plotted for three different levels along with the DRNL output **(A)** and the linear path component of the DRNL model **(B)** for all three vocalizations. Black lines correspond to the PSTHs and gray shaded areas correspond to the DRNL outputs. The best frequency and the peak correlation value for the recording sites are labeled on the top-right of the plots.

we cannot claim that these simplified manipulations will exactly replicate the stimulus envelope at each location along the basilar membrane for all stimuli and across different guinea pigs. Additionally, we had to estimate the BF of each ICC site to then calculate the different BF-dependent parameters for the DRNL model. Our method of BF calculation may not always estimate the true BF for each ICC site. Therefore, these potential biases could have contributed to the wide range of patterns observed in this study, and thus more ICC cases may have actually been capable of following the envelope of the vocalizations than we presented in our results.

Although anesthesia has shown to cause changes in spontaneous activity as well as acoustic-driven temporal and spectral patterns in the auditory cortex, there seem to be minimal changes that occur within the IC, at least when using ketamine (Zurita et al., 1994; Astl et al., 1996; Gaese and Ostwald, 2001; Suta et al., 2003; Syka et al., 2005; Ter-Mikaelian et al., 2007). One study used a combination of ketamine and sodium pentobarbital in gerbils and reported no or minimal changes compared to the awake state for spontaneous activity. In response to artificial stimuli (i.e., pure tones, noise, amplitude modulated sinusoids), they also observed no or minimal changes for first spike latency, variation in first spike latency, trial-by-trial reliability, temporal synchronization capabilities, and acoustic-driven firing rate. Similar to our study, several studies used a combination of ketamine and xylazine in guinea pigs and there appeared to be no or minimal changes compared to the awake state for spontaneous activity, firing patterns in response to pure tones, and tuning properties based on Q_{10} values (Astl et al., 1996; Torterolo et al., 2002; Suta et al., 2003). Although a few studies have shown some changes in spontaneous activity and acoustic-driven response patterns (e.g., spike rate, thresholds, latencies) in the IC due to anesthesia (Kuwada et al., 1989; Astl et al., 1996; Torterolo et al., 2002; Tollin et al., 2004), those effects were caused by barbituates that may have altered spiking properties more dramatically than ketamine. Therefore, the use of ketamine in our study is not expected to have dramatically altered the temporal following capabilities of ICC neurons in response to vocalizations. Even if there were some anesthetic effects on our ICC responses, it would seem that the following capabilities would become worse with anesthesia, and thus more ICC neurons may have exhibited high correlation values if our study was performed in awake animals. It seems unlikely that ketamine would somehow improve the following capabilities of ICC neurons to natural stimuli that was not originally possible in the awake state.

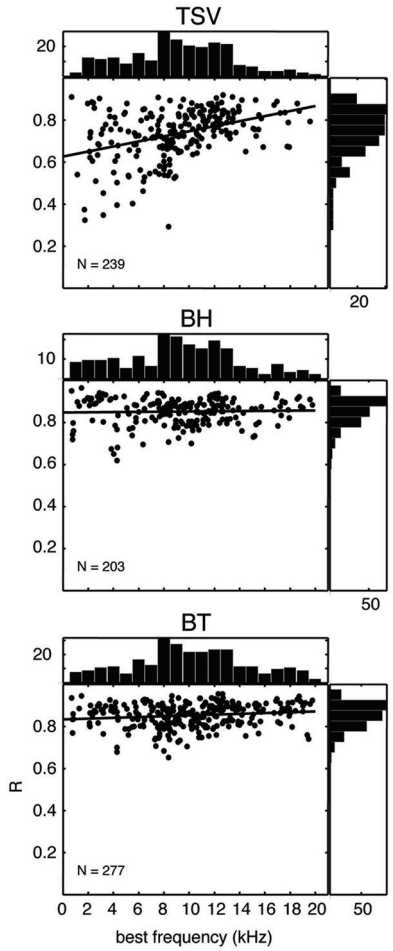

FIGURE 10 | Scatter plots for correlation values versus best frequency. The peak correlation values (*R*) is plotted for ICC sites with different best frequencies for all three vocalizations from 10 animals. The histogram on top of each plot shows the distribution of best frequencies across all the ICC sites, while the histogram on the right shows the distribution of correlation values. The solid line in each plot corresponds to the least-squares fit to the data. *N* denotes number of total ICC sites for each vocalization.

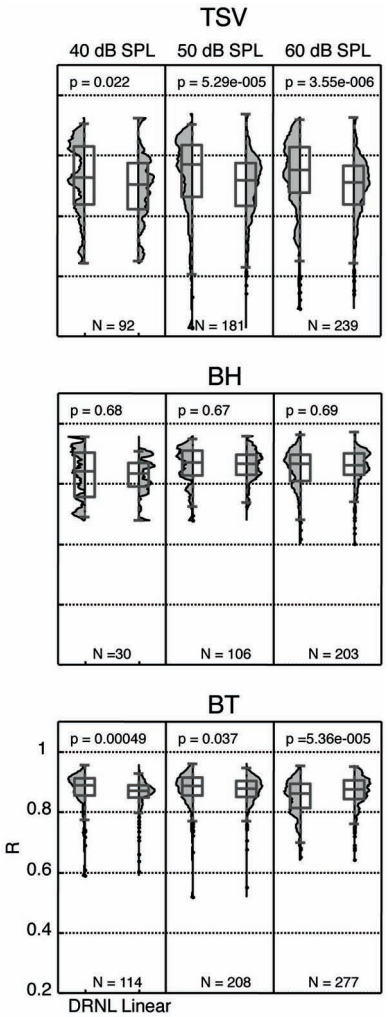

FIGURE 11 | Distribution of correlation values for different levels. The peak correlation values (*R*) are plotted for both the linear path and DRNL outputs for three different stimulus levels for each vocalization from 10 animals. The *R* values were significantly different between the linear path output and the DRNL output for some cases ($p < 0.05$, Wilcoxon rank-sum test). *N* denotes number of total ICC sites for each condition.

We analyzed multi-unit activity instead of single-unit responses. Previous studies have shown that individual ICC neurons have slow membrane time constants (e.g., 12–33 ms) and do not typically fire more than once within a 10 ms window (Sivaramakrishnan and Oliver, 2006; Chen et al., 2012). It is possible that the PSTH response to individual ICC neurons may not be able to fully follow the BF-matched envelope of different vocalizations because they cannot continuously keep up with the entire pattern over time. However, multi-unit activity consisting of clusters of neurons surrounding a given ICC site may exhibit greater following capabilities because each neuron could fill in different parts of the stimulus envelope. As a result, the correlation values and examples presented in our study may be higher than what are typically observed for individual neurons. Regardless, some individual ICC neurons must still be able to follow the stimulus envelope, even if just certain portions of the envelope, so that together with other neurons they can elicit the types of patterns observed in our study. It will be interesting to investigate this possibility and if different neurons located along the same ICC lamina are designed to accurately represent the envelope structure of natural stimuli as a temporal code across a cluster or population of neurons. At least for AMI implementation, electrical stimulation of a given site will activate a cluster of neurons surrounding that site. Therefore, AMI stimulation may activate enough neurons surrounding each site to still achieve the envelope following capabilities that were observed for a subset of our multi-unit recordings.

PARALLEL SUB-LEMNISCAL CODING PATHWAYS

Anatomical studies have shown that spatially distinct functional zones exist within the ICC that receive different combinations of inputs from lower brainstem centers (Roth et al., 1978; Brunso-Bechtold et al., 1981; Shneiderman and Henkel, 1987; Oliver

et al., 1997; Cant and Benson, 2003, 2006; Loftus et al., 2004). Furthermore, electrophysiological studies have demonstrated that different regions along the isofrequency ICC laminae process sound information (e.g., threshold, frequency tuning, frequency sweep speed, best modulation frequency, latency, binaurality) in different ways (Semple and Aitkin, 1979; Stiebler, 1986; Langner and Schreiner, 1987; Schreiner and Langner, 1988; Ehret, 1997; Hage and Ehret, 2003). Therefore, it is expected that different ICC neurons would exhibit varying spiking patterns to the same sound stimulus.

What is particularly interesting about our results is that some ICC neurons could closely follow the envelope of the vocalizations while other neurons exhibited complex and dissimilar spiking patterns. Based on the studies described above, it is not surprising that we observed these differences in spiking patterns across different ICC neurons. The question arises as to whether this subset of neurons that can closely follow the envelope of natural stimuli are scattered throughout the ICC or if these neurons are located in a specific region within the ICC. The latter scenario would be advantageous for the AMI in which the electrode array could be implanted within a specific ICC region to potentially transmit sufficient envelope cues to higher perceptual centers. In our study, we did not histologically confirm the location of our recording sites throughout the ICC to answer that question. However, there are several anatomical and physiological studies described below that have revealed a specific ICC region that exhibits more robust temporal (and spectral) coding properties than the other ICC regions, and thus may correspond to the location of the neurons that closely followed the envelope of the vocalizations in our study.

In gerbils, it was shown that neurons from the cochlear nuclei and nuclei of the lateral lemniscus project throughout the ICC, while inputs from the superior olivary complex project predominantly to more rostral and lateral locations of the ICC (Cant and Benson, 2006). These anatomical results suggested that at least two distinct regions exist within the ICC: a caudal–medial region and a more rostral–lateral region. The authors further showed that these same regions project to distinct locations along the caudal–rostral (isofrequency) dimension of the ventral division of the medial geniculate nucleus (MGv): the caudal–medial ICC region projects to the caudal third of the MGv while the rostral–lateral ICC region projects to the rostral two-thirds of the MGv (Cant and Benson, 2007). Based on anatomical studies in cats and rats (Rodrigues-Dagaeff et al., 1989; Polley et al., 2007; Storace et al., 2010), the rostral MGv then projects throughout the auditory cortex, including primary auditory cortex (A1), while the caudal MGv projects predominantly to regions outside of A1. Consistent with the spatial organization shown in the anatomical studies and based on physiological studies in cats and guinea pigs, the pathway through more rostral–lateral ICC regions (versus caudal–medial regions along an isofrequency lamina) and more rostral MGv regions (versus caudal regions along an isofrequency lamina) up to A1 exhibits neural activity with lower thresholds, more excitatory activity, greater spatial synchrony across neurons, shorter latencies, and less temporal jitter of spiking to a given stimulus (Rodrigues-Dagaeff et al., 1989; Lim and Anderson, 2007; Straka et al., 2013). This pathway through more rostral MGv regions has also shown to exhibit more precise tonotopy, sharper frequency tuning, and more time-locked spiking to repetitive clicks (Rodrigues-Dagaeff et al., 1989). Together, these findings across different species suggest that there are neurons located in a specific ICC region (i.e., a rostral–lateral location along the isofrequency laminae) that project to rostral MGv up to A1 and are designed to robustly code for sound stimuli. Therefore, the ICC responses that most accurately matched the envelope of the different vocalizations in our study may correspond to neurons within this "rostral–lateral" pathway.

IMPLICATIONS FOR AMI IMPLEMENTATION

Speech understanding in quiet environments, which is the initial goal for the AMI, is possible with envelope modulations as low as 50 Hz (Shannon et al., 1995; Nie et al., 2006; Xu and Pfingst, 2008). Assuming that the rostral–lateral pathway described in the previous section exists, especially within the human ICC, CI-based strategies may still be effective for the AMI. In particular, if it is possible to target the subset of ICC neurons that can robustly code for the stimulus envelope (up to ∼100 Hz), then it may be possible to restore speech perception by stimulating each frequency lamina of the ICC with amplitude-modulated pulse trains matching the corresponding bandpass-filtered envelope of the stimulus based on a DRNL model fitted for humans (Lopez-Poveda and Meddis, 2001).

There are several possibilities as to why the AMI has not yet been able to restore sufficient speech understanding even though a CI-based strategy is being used in the implanted patients. Stimulation of just a single site within each lamina may not activate enough neurons to transmit sufficient envelope cues to higher centers. An alternative possibility is that the current patients were not implanted in the correct region or there is a lack of a specific region to access the subset of ICC neurons identified in our study that could temporally code for the stimulus envelope. Based on the studies described in the previous section, there appears to be a rostral–lateral region along the ICC laminae that could correspond to the location of this subset of robustly coding ICC neurons. However, we need to perform additional studies with histological reconstructions to confirm that the neurons with the highest correlation values across different vocalizations truly correspond to that rostral–lateral region. Furthermore, we observed complex neural patterns across different levels for many ICC neurons. Although we observed some neurons that exhibited high correlation values across multiple levels, we need to investigate if the DNRL model can be expanded to better predict the level-dependent effects across a greater number of ICC neurons since good speech perception may not be achieved with activation of just a few neurons.

More recently, Calixto et al. (2012) has shown that electrical stimulation of a single site in a given isofrequency lamina artificially induces strong inhibitory and suppressive effects within the auditory cortex. However, co-activation of at least two regions within a lamina is able to overcome much of these inhibitory and suppressive properties. The AMI may still be able to use a CI-based strategy to activate the ICC neurons that are capable of following the stimulus envelope, but by stimulating at least two

sites within each lamina. Implanting a two-shank electrode array into the ICC to position two sites within each lamina is surgically feasible. Currently, we are investigating and developing a two-shank array device to safely implant and implement in future AMI patients.

ACKNOWLEDGMENTS

This work was supported by the German Ministry of Research and Education (grant 01GQ0816), Cochlear Limited, SFB 599, start-up funds from the University of Minnesota (Institute for Translational Neuroscience and the College of Science and Engineering), and German Research Foundation (PP 1608 Kr 3370/2-1). We would like to thank Roger Calixto, Anke Neuheiser, Henning Voigt, David Edge, Michael Herrmann, Dmitri Bibitchchkov, and Dominika Lyzwa for assistance and suggestions during the experiments and data analysis. We would also like to thank Jasmine Grimsley for her suggestions in how to record and process the vocalizations from guinea pigs and Simon Doclo for his guidance with the adaptive filtering used for sound calibration. We appreciate the generosity of Ray Meddis, Christian Sumner, and Mark Steadman for providing the middle ear-DRNL Matlab code and helping us with its implementation for our analysis.

REFERENCES

Astl, J., Popelar, J., Kvasnak, E., and Syka, J. (1996). Comparison of response properties of neurons in the inferior colliculus of guinea pigs under different anesthetics. *Audiology* 35, 335–345. doi: 10.3109/00206099609071954

Bledsoe, S. C., Shore, S. E., and Guitton, M. J. (2003). Spatial representation of corticofugal input in the inferior colliculus: a multicontact silicon probe approach. *Exp. Brain Res.* 153, 530–542. doi: 10.1007/s00221-003-1671-6

Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *J. Acoust. Soc. Am.* 113, 369–388. doi: 10.1121/1.1519544

Brunso-Bechtold, J. K., Thompson, G. C., and Masterton, R. B. (1981). HRP study of the organization of auditory afferents ascending to central nucleus of inferior colliculus in cat. *J. Comp. Neurol.* 197, 705–722. doi: 10.1002/cne.901970410

Calixto, R., Lenarz, M., Neuheiser, A., Scheper, V., Lenarz, T., and Lim, H. H. (2012). Coactivation of different neurons within an isofrequency lamina of the inferior colliculus elicits enhanced auditory cortical activation. *J. Neurophysiol.* 108, 1199–1210. doi: 10.1152/jn.00111.2012

Cant, N. B., and Benson, C. G. (2003). Parallel auditory pathways: projection patterns of the different neuronal populations in the dorsal and ventral cochlear nuclei. *Brain Res. Bull.* 60, 457–474. doi: 10.1016/S0361-9230(03)00050-9

Cant, N. B., and Benson, C. G. (2006). Organization of the inferior colliculus of the gerbil (*Meriones unguiculatus*): differences in distribution of projections from the cochlear nuclei and the superior olivary complex. *J. Comp. Neurol.* 495, 511–528. doi: 10.1002/cne.20888

Cant, N. B., and Benson, C. G. (2007). Multiple topographically organized projections connect the central nucleus of the inferior colliculus to the ventral division of the medial geniculate nucleus in the gerbil, *Meriones unguiculatus*. *J. Comp. Neurol.* 503, 432–453. doi: 10.1002/cne.21391

Chen, C., Read, H. L., and Escabi, M. A. (2012). Precise feature based time scales and frequency decorrelation lead to a sparse auditory code. *J. Neurosci.* 32, 8454–8468. doi: 10.1523/JNEUROSCI.6506-11.2012

Cooper, N. P., and Yates, G. K. (1994). Nonlinear input–output functions derived from the responses of guinea-pig cochlear nerve fibres: variations with characteristic frequency. *Hear. Res.* 78, 221–234. doi: 10.1016/0378-5955(94)90028-0

Dugue, P., Le Bouquin-Jeannes, R., Edeline, J. M., and Faucon, G. (2010). A physiologically based model for temporal envelope encoding in human primary auditory cortex. *Hear. Res.* 268, 133–144. doi: 10.1016/j.heares.2010.05.014

Ehret, G. (1997). "The auditory midbrain, a "shunting yard" of acoustical information processing," in *The Central Auditory System*, eds G. Ehret and R. Romand (New York: Oxford University Press, Inc.), 259–316.

Evans, E. F. (1972). The frequency response and other properties of single fibres in the guinea-pig cochlear nerve. *J. Physiol.* 226, 263–287.

Frisina, R. D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. *Hear. Res.* 158, 1–27. doi: 10.1016/S0378-5955(01)00296-9

Gaese, B. H., and Ostwald, J. (2001). Anesthesia changes frequency tuning of neurons in the rat primary auditory cortex. *J. Neurophysiol.* 86, 1062–1066.

Geniec, P., and Morest, D. K. (1971). The neuronal architecture of the human posterior colliculus. A study with the Golgi method. *Acta Otolaryngol. Suppl.* 295, 1–33.

Gourevitch, B., Doisy, T., Avillac, M., and Edeline, J. M. (2009). Follow-up of latency and threshold shifts of auditory brainstem responses after single and interrupted acoustic trauma in guinea pig. *Brain Res.* 1304, 66–79. doi: 10.1016/j.brainres.2009.09.041

Grimsley, J. M., Shanbhag, S. J., Palmer, A. R., and Wallace, M. N. (2012). Processing of communication calls in Guinea pig auditory cortex. *PLoS ONE* 7:e51646. doi: 10.1371/journal.pone.0051646

Guerin, A., Jeannes Rle, B., Bes, J., Faucon, G., and Lorenzi, C. (2006). Evaluation of two computational models of amplitude modulation coding in the inferior colliculus. *Hear. Res.* 211, 54–62. doi: 10.1016/j.heares.2005.10.001

Hage, S. R., and Ehret, G. (2003). Mapping responses to frequency sweeps and tones in the inferior colliculus of house mice. *Eur. J. Neurosci.* 18, 2301–2312. doi: 10.1046/j.1460-9568.2003.02945.x

Harper, L. (1976). "Behavior," in *The Biology of the Guinea Pig*, eds J. Wagner and P. Manning (New York: Academic Press), 31–51.

Hewitt, M. J., and Meddis, R. (1994). A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus. *J. Acoust. Soc. Am.* 95, 2145–2159. doi: 10.1121/1.408676

Holmes, S. D., Sumner, C. J., O'Mard, L. P., and Meddis, R. (2004). The temporal representation of speech in a nonlinear model of the guinea pig cochlea. *J. Acoust. Soc. Am.* 116, 3534–3545. doi: 10.1121/1.1815111

Holmstrom, L. A., Eeuwes, L. B., Roberts, P. D., and Portfors, C. V. (2010). Efficient encoding of vocalizations in the auditory midbrain. *J. Neurosci.* 30, 802–819. doi: 10.1523/JNEUROSCI.1964-09.2010

Huetz, C., Gourevitch, B., and Edeline, J. M. (2011). Neural codes in the thalamocortical auditory system: from artificial stimuli to communication sounds. *Hear. Res.* 271, 147–158. doi: 10.1016/j.heares.2010.01.010

Irino, T., and Patterson, R. D. (2006). A dynamic compressive gammachirp auditory filterbank. *IEEE Trans. Audio Speech Lang. Process.* 14, 2222–2232. doi: 10.1109/TASL.2006.874669

Jepsen, M. L., Ewert, S. D., and Dau, T. (2008). A computational model of human auditory signal processing and perception. *J. Acoust. Soc. Am.* 124, 422–438. doi: 10.1121/1.2924135

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J. Acoust. Soc. Am.* 68, 1115–1122. doi: 10.1121/1.384982

Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577. doi: 10.1152/physrev.00029.2003

Krishna, B. S., and Semple, M. N. (2000). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J. Neurophysiol.* 84, 255–273.

Kuwada, S., Batra, R., and Stanford, T. R. (1989). Monaural and binaural response properties of neurons in the inferior colliculus of the rabbit: effects of sodium pentobarbital. *J. Neurophysiol.* 61, 269–282.

Langner, G., Albert, M., and Briede, T. (2002). Temporal and spatial coding of periodicity information in the inferior colliculus of awake chinchilla (*Chinchilla laniger*). *Hear. Res.* 168, 110–130. doi: 10.1016/S0378-5955(02)00367-2

Langner, G., and Schreiner, C. E. (1987). "Topology of functional parameters in the inferior colliculus of the cat," in *New Frontiers in Brain Research*, eds N. Elsner and O. D. Creutzfeldt (Stuttgart: Thieme), 122.

Langner, G., and Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.* 60, 1799–1822.

Lenarz, M., Lim, H. H., Patrick, J. F., Anderson, D. J., and Lenarz, T. (2006). Electrophysiological validation of a human prototype auditory

midbrain implant in a guinea pig model. *J. Assoc. Res. Otolaryngol.* 7, 383–398. doi: 10.1007/s10162-006-0056-5

Lim, H. H., and Anderson, D. J. (2006). Auditory cortical responses to electrical stimulation of the inferior colliculus: implications for an auditory midbrain implant. *J. Neurophysiol.* 96, 975–988. doi: 10.1152/jn.01112.2005

Lim, H. H., and Anderson, D. J. (2007). Spatially distinct functional output regions within the central nucleus of the inferior colliculus: implications for an auditory midbrain implant. *J. Neurosci.* 27, 8733–8743. doi: 10.1523/JNEUROSCI.5127-06.2007

Lim, H. H., Lenarz, T., Joseph, G., Battmer, R. D., Samii, A., Samii, M., et al. (2007). Electrical stimulation of the midbrain for hearing restoration: insight into the functional organization of the human central auditory system. *J. Neurosci.* 27, 13541–13551. doi: 10.1523/JNEUROSCI.3123-07.2007

Lim, H. H., Lenarz, M., Joseph, G., and Lenarz, T. (2013). Frequency representation within the human brain: stability versus plasticity. *Sci. Rep.* 3, 1474. doi: 10.1038/srep01474

Lim, H. H., Lenarz, M., and Lenarz, T. (2009). Auditory midbrain implant: a review. *Trends Amplif.* 13, 149–180. doi: 10.1177/1084713809348372

Loftus, W. C., Bishop, D. C., Saint Marie, R. L., and Oliver, D. L. (2004). Organization of binaural excitatory and inhibitory inputs to the inferior colliculus from the superior olive. *J. Comp. Neurol.* 472, 330–344. doi: 10.1002/cne.20070

Lopez-Poveda, E. A., and Meddis, R. (2001). A human nonlinear cochlear filterbank. *J. Acoust. Soc. Am.* 110, 3107–3118. doi: 10.1121/1.1416197

Malmierca, M. S., Rees, A., Le Beau, F. E., and Bjaalie, J. G. (1995). Laminar organization of frequency-defined local axons within and between the inferior colliculi of the guinea pig. *J. Comp. Neurol.* 357, 124–144. doi: 10.1002/cne.903570112

McKay, C. M., Lim, H. H., and Lenarz, T. (2013). Temporal processing in the auditory system: insights from cochlear and auditory midbrain implantees. *J. Assoc. Res. Otolaryngol.* 14, 103–124. doi: 10.1007/s10162-012-0354-z

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (2001). A computational algorithm for computing nonlinear auditory frequency selectivity. *J. Acoust. Soc. Am.* 109, 2852–2861. doi: 10.1121/1.1370357

Moonen, M., and Proudler, I. (1998). *An Introduction to Adaptive Signal Processing.* Available at: http://homes.esat.kuleuven.be/~moonen/asp_course.html [accessed March, 2010].

Nelson, P. C., and Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J. Acoust. Soc. Am.* 116, 2173–2186. doi: 10.1121/1.1784442

Neuheiser, A., Lenarz, M., Reuter, G., Calixto, R., Nolte, I., Lenarz, T., et al. (2010). Effects of pulse phase duration and location of stimulation within the inferior colliculus on auditory cortical evoked potentials in a guinea pig model. *J. Assoc. Res. Otolaryngol.* 11, 689–708. doi: 10.1007/s10162-010-0229-0

Nie, K., Barco, A., and Zeng, F. G. (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear Hear.* 27, 208–217. doi: 10.1097/01.aud.0000202312.31837.25

Oliver, D. L. (2005). "Neuronal organization in the inferior colliculus," in *The Inferior Colliculus*, eds J. A. Winer and C. E. Schreiner (New York: Springer Science + Business Media), 69–114.

Oliver, D. L., Beckius, G. E., Bishop, D. C., and Kuwada, S. (1997). Simultaneous anterograde labeling of axonal layers from lateral superior olive and dorsal cochlear nucleus in the inferior colliculus of cat. *J. Comp. Neurol.* 382, 215–229. doi: 10.1002/(SICI)1096-9861(19970602)382:2<215::AID-CNE6>3.0.CO;2-6

Philibert, B., Laudanski, J., and Edeline, J. M. (2005). Auditory thalamus responses to guinea-pig vocalizations: a comparison between rat and guinea-pig. *Hear. Res.* 209, 97–103. doi: 10.1016/j.heares.2005.07.004

Polley, D. B., Read, H. L., Storace, D. A., and Merzenich, M. M. (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *J. Neurophysiol.* 97, 3621–3638. doi: 10.1152/jn.01298.2006

Rees, A., and Moller, A. R. (1987). Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear. Res.* 27, 129–143. doi: 10.1016/0378-5955(87)90014-1

Rees, A., and Palmer, A. R. (1989). Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J. Acoust. Soc. Am.* 85, 1978–1994. doi: 10.1121/1.397851

Rees, A., Sarbaz, A., Malmierca, M. S., and Le Beau, F. E. (1997). Regularity of firing of neurons in the inferior colliculus. *J. Neurophysiol.* 77, 2945–2965.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer technique. *J. Acoust. Soc. Am.* 49(Suppl. 2), 1218–1231. doi: 10.1121/1.1912485

Rodrigues-Dagaeff, C., Simm, G., De Ribaupierre, Y., Villa, A., De Ribaupierre, F., and Rouiller, E. M. (1989). Functional organization of the ventral division of the medial geniculate body of the cat: evidence for a rostro-caudal gradient of response properties and cortical projections. *Hear. Res.* 39, 103–125. doi: 10.1016/0378-5955(89)90085-3

Roth, G. L., Aitkin, L. M., Andersen, R. A., and Merzenich, M. M. (1978). Some features of the spatial organization of the central nucleus of the inferior colliculus of the cat. *J. Comp. Neurol.* 182, 661–680. doi: 10.1002/cne.901820407

Schreiner, C. E., and Langner, G. (1988). Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J. Neurophysiol.* 60, 1823–1840.

Semple, M. N., and Aitkin, L. M. (1979). Representation of sound frequency and laterality by units in central nucleus of cat inferior colliculus. *J. Neurophysiol.* 42, 1626–1639.

Shannon, R. V., Fu, Q. J., and Galvin, J. III. (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol. Suppl.* 552, 50–54. doi: 10.1080/03655230410017562

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303

Shneiderman, A., and Henkel, C. K. (1987). Banding of lateral superior olivary nucleus afferents in the inferior colliculus: a possible substrate for sensory integration. *J. Comp. Neurol.* 266, 519–534. doi: 10.1002/cne.902660406

Sivaramakrishnan, S., and Oliver, D. L. (2006). Neuronal responses to lemniscal stimulation in laminar brain slices of the inferior colliculus. *J. Assoc. Res. Otolaryngol.* 7, 1–14. doi: 10.1007/s10162-005-0017-4

Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 416, 87–90. doi: 10.1038/416087a

Snyder, R. L., Bierer, J. A., and Middlebrooks, J. C. (2004). Topographic spread of inferior colliculus activation in response to acoustic and intracochlear electric stimulation. *J. Assoc. Res. Otolaryngol.* 5, 305–322. doi: 10.1007/s10162-004-4026-5

Snyder, R. L., Vollmer, M., Moore, C. M., Rebscher, S. J., Leake, P. A., and Beitel, R. E. (2000). Responses of inferior colliculus neurons to amplitude-modulated intracochlear electrical pulses in deaf cats. *J. Neurophysiol.* 84, 166–183.

Stiebler, I. (1986). Tone-threshold mapping in the inferior colliculus of the house mouse. *Neurosci. Lett.* 65, 336–340. doi: 10.1016/0304-3940(86)90285-5

Storace, D. A., Higgins, N. C., and Read, H. L. (2010). Thalamic label patterns suggest primary and ventral auditory fields are distinct core regions. *J. Comp. Neurol.* 518, 1630–1646. doi: 10.1002/cne.22345

Straka, M., Schmitz, S. J., Lee, P. T., Mcmahon, M., and Lim, H. H. (2013). Responses to acoustic stimuli vary along the isofrequency laminae of the inferior colliculus consistent with spatially segregated sublemniscal auditory pathways. *Assoc. Res. Otolaryngol. Abstr.* 36, 324.

Sumner, C. J., O'Mard, L. P., Lopez-Poveda, E. A., and Meddis, R. (2003). A nonlinear filter-bank model of the guinea-pig cochlear nerve: rate responses. *J. Acoust. Soc. Am.* 113, 3264–3274. doi: 10.1121/1.1568946

Suta, D., Kvasnak, E., Popelar, J., and Syka, J. (2003). Representation of species-specific vocalizations in the inferior colliculus of the guinea pig. *J. Neurophysiol.* 90, 3794–3808. doi: 10.1152/jn.01175.2002

Syka, J., Popelar, J., Kvasnak, E., and Astl, J. (2000). Response properties of neurons in the central nucleus and external and dorsal cortices of the inferior colliculus in guinea pig. *Exp. Brain Res.* 133, 254–266. doi: 10.1007/s002210000426

Syka, J., Suta, D., and Popelar, J. (2005). Responses to species-specific vocalizations in the auditory cortex of awake and anesthetized guinea pigs. *Hear. Res.* 206, 177–184. doi: 10.1016/j.heares.2005.01.013

Ter-Mikaelian, M., Sanes, D. H., and Semple, M. N. (2007). Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. *J. Neurosci.* 27, 6091–6102. doi: 10.1523/JNEUROSCI.4848-06.2007

Tollin, D. J., Populin, L. C., and Yin, T. C. (2004). Neural correlates of the precedence effect in the inferior colliculus of behaving cats. *J. Neurophysiol.* 92, 3286–3297. doi: 10.1152/jn.00606.2004

Torterolo, P., Falconi, A., Morales-Cobas, G., and Velluti, R. A. (2002). Inferior colliculus unitary activity in wakefulness, sleep and under barbiturates. *Brain Res.* 935, 9–15. doi: 10.1016/S0006-8993(02)02235-7

Wang, X., Lu, T., Bendor, D., and Bartlett, E. (2008). Neural coding of temporal information in auditory thalamus and cortex. *Neuroscience* 154, 294–303. doi: 10.1016/j.neuroscience.2008.03.065

Woolley, S. M., Gill, P. R., and Theunissen, F. E. (2006). Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *J. Neurosci.* 26, 2499–2512. doi: 10.1523/JNEUROSCI.3731-05.2006

Xu, L., and Pfingst, B. E. (2008). Spectral and temporal cues for speech recognition: implications for auditory prostheses. *Hear. Res.* 242, 132–140. doi: 10.1016/j.heares.2007.12.010

Yates, G. K., Winter, I. M., and Robertson, D. (1990). Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range. *Hear. Res.* 45, 203–219. doi: 10.1016/0378-5955(90)90121-5

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (2001). A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.* 109, 648–670. doi: 10.1121/1.1336503

Zheng, Y., and Escabi, M. A. (2008). Distinct roles for onset and sustained activity in the neuronal code for temporal periodicity and acoustic envelope shape. *J. Neurosci.* 28, 14230–14244. doi: 10.1523/JNEUROSCI.2882-08.2008

Zilany, M. S., Bruce, I. C., Nelson, P. C., and Carney, L. H. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *J. Acoust. Soc. Am.* 126, 2390–2412. doi: 10.1121/1.3238250

Zurita, P., Villa, A. E., De Ribaupierre, Y., De Ribaupierre, F., and Rouiller, E. M. (1994). Changes of single unit activity in the cat's auditory thalamus and cortex associated to different anesthetic conditions. *Neurosci. Res.* 19, 303–316. doi: 10.1016/0168-0102(94)90043-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 March 2013; accepted: 24 September 2013; published online: 21 October 2013.

Citation: Rode T, Hartmann T, Hubka P, Scheper V, Lenarz M, Lenarz T, Kral A and Lim HH (2013) Neural representation in the auditory midbrain of the envelope of vocalizations based on a peripheral ear model. Front. Neural Circuits ***7****:166. doi: 10.3389/fncir.2013.00166*

This article was submitted to the journal Frontiers in Neural Circuits.

Copyright © 2013 Rode, Hartmann, Hubka, Scheper, Lenarz, Lenarz, Kral and Lim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.