Identifying the Cellular Sources of the Low-Frequency Cochlear Response

By

Aryn M Kamerer
B.A., University of Kansas, 2010

Submitted to the graduate degree program in Communication Sciences and Disorders and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Mark Chertoff, PhD

John Ferraro, PhD

Tiffany Johnson, PhD

Jon Brumberg, PhD

Francisco J Diaz, PhD

Date Defended: 24 August 2017

The dissertation committee for Aryn M Kamerer certifies that this is the approved version of the following dissertation:

Identifying the Cellular Sources of the Low-Frequency Cochlear Response

Chair: Mark Chertoff, PhD

Date Approved: 29 August 2017

ABSTRACT

Medical and technologic treatments for hearing loss are quickly outpacing current clinical diagnostic techniques. In order to knowledgeably treat patients and accurately predict treatment outcomes, we need diagnostic tools that can identify the anatomic damage or dysfunction underlying the loss of hearing. The cochlear microphonic –a reflection of current flow through outer hair cells –in conjunction with high-pass noise or suppressing tones, shows promise as a method of assessing the health of outer hair cells at specific locations along the cochlear partition in rodent models. In this study, we propose that the electrical potential recorded from the round window in gerbils to *low-frequency* tones contains responses from a number of cellular sources in addition to and, in some high-pass noise conditions, more substantial than the outer hair cell response. A model is created from data found in existing literature and our previous studies, which provides evidence for identifying each cellular source contributing to this low-frequency round window potential, termed the cochlear response (CR).

The CR was recorded via an electrode placed in the round window niche of 16 Mongolian gerbils and elicited with a 45 Hz tone burst embedded in 18 high-pass filtered noise conditions, in order to target responses from increasing apical-to-basal regions along the cochlear partition. Independent component analysis and filtering of the response recovered several contributing sources whose spectral and temporal content provided clues to their identification. These suspected sources were modeled using previously-published hair cell and auditory nerve response data, and then weighted and combined using linear regression to produce a model response that fits closely to the mean CR waveform.

We conclude that the low-frequency CR contains contributions from several cellular sources, for which the model provides evidence of the most significant contributors being outer hair cells, inner hair cells, and apical auditory nerve fibers. Therefore, the CR shows the capacity to be developed into a diagnostic tool that can assess the health of multiple structures in the cochlea objectively, simultaneously, and independently.

Supported by NIH R33DC011096 and the University of Kansas Office of Graduate Studies

ACKNOWLEDGEMENTS

I would never have been able to finish my dissertation without the guidance of my mentors and committee members, the support of my friends and family, and of course, money from the government.

I want to express the deepest gratitude toward my mentor, Mark Chertoff. Dr. Chertoff's easy-going attitude toward life and serious enthusiasm for science made me feel at home from my first day in the lab. Over the last five years he has taught me to question everyone and everything, including myself. Luckily, after skating past the brink of identity crisis, I truly believe I am a better scientist and a more whole human being today. Thank you.

I am also grateful to all of those with whom I have had the pleasure to work during this and other projects. Each member of my dissertation committee has provided me professional and personal guidance and taught me much about scientific research, teaching, and life. I appreciate your time and efforts over the years.

Nobody has been more important to me in the pursuit of this project than my family. I would like to thank my parents who have always supported my wildest academic dreams and given me every opportunity to fulfil my career and personal goals. Thank you to my brothers who are all brilliant and continuously celebrate my successes. Thank you to all my friends over the years and to Spencer, whose adventurous spirit and support kept me afloat and looking toward the future in the final stages of my project.

And last, but certainly not least, I want to give a special thanks to my grandfather, Benjamin Friesen, whom I admire in all facets of my life. He is a true scientist, and through example has taught me to have an open mind and to think logically and be inquisitive, as well as instilled in me the importance of creativity and a love of the arts. There is no one in this world whom I aspire to be more.

TABLE OF CONTENTS

INTRODUCTION

Hearing loss is one of the most prevalent disorders in the United States and across the globe. Clinical audiologists are often the primary health professionals who diagnose and treat these hearing disorders in the US. In order to diagnose a hearing loss, clinical audiologists perform a battery of tests that help the clinician characterize the hearing loss and locate the site-of-lesion, both of which are crucial to choose an effective treatment for their patients' individual needs. In a somewhat recent revelation, the field of hearing science has decided that the audiogram –the diagnostic tool on which clinicians rely most heavily –is not adequate at determining the site-of-lesion of a sensorineural hearing loss, that is, a hearing loss generated from damage or dysfunction to the cochlea or central auditory system (Landegger, Psaltis, & Stankovic, 2016). In the cochlea, specifically, there are many anatomical structures that if damaged, will result in a hearing loss. Generally, hearing aids are prescribed to treat most sensorineural hearing losses, but hearing aid outcomes are highly variable even between patients with similar hearing loss patterns. Since there is no effective umbrella treatment for all types of sensorineural hearing loss, it is imperative that diagnostic measures differentiate the possible sites-of-lesion within the cochlea, both spatially and cell-specifically. More sensitive diagnostic measures could spur the development of differentiated hearing aid technology, more effective cochlear implantation, and research into new treatment methods. More immediately, they could at least provide some predictability of a given patient's treatment outcome with hearing aids.

Objective measures of cochlear health are the most promising tools for identifying damage or physiological dysfunction underlying a given hearing loss. The auditory brainstem response (ABR) and otoacoustic emissions (OAEs) are widely used to screen for retrocochlear and outer hair cell (OHC) damage, respectively, although presence and threshold of these responses are the only parameters used clinically. Electrocochleography is a more specific measure of the peripheral auditory nerve (AN) response, where certain parameters of the compound action potential (CAP) can be used to diagnose Meniere's disease; and a lot of research in the last decade focused on measurements of amplitude and morphology that may be used to assess specific disorders within the AN. Another auditory-evoked

potential is the cochlear microphonic (CM), which is a reflection of current flow through OHCs and therefore may be developed into a diagnostic measure (Mary Ann Cheatham, Naik, & Dallos, 2011; Withnell, 2001). Research by Chertoff and colleagues (2012, 2014, 2015), Kamerer et al. (2016) and Charaziak et al. (2017) do in fact provide evidence in rodent models for utilizing the CM to measure OHCs for location-specific assessment.

The CM reported in these studies are recorded with the primary electrode resting in the round window (RW) niche of the cochlea. Recording from the RW could pose a problem, as a high-frequency stimulus will cause maximal displacement of the basilar membrane (BM) in the base, proximal to the electrode, resulting in large phase incoherence and causing a reduction in the response amplitude due to cancellations of the response wave (P Dallos, 1973; Whitfield & Ross, 1965). This could lead to false diagnoses of damage. A solution to this is to lower the stimulus frequency, however, this leads to a second problem: the responses from all the OHCs in the cochlea will be weighted by their Euclidean distance to the electrode, meaning OHCs residing in the base will overpower responses from OHCs in the apex. This could confound attempts to diagnose the health of OHCs in the apex (Dallos, 1969; Patuzzi et al., 1989). Several studies have developed methods of solving this issue, using a secondary noise to suppress, or mask –a term which will be used in this manner throughout this report (see *Appendix A: Cochlear Microphonic* for definitions and uses of these terms)–responses from OHCs outside a targeted region of the cochlear partition.

Chertoff et al. (2012, 2014) used low-frequency stimulus tones embedded in a high-pass noise masking paradigm of seventeen consecutively-increasing cutoff frequencies to record the CM in gerbils from increasing regions of the cochlea. Chertoff and colleagues converted the cutoff frequency of each masking noise condition to a location along the cochlear partition and then plotted the amplitude of ~750 Hz tone burst-elicited CM at each location. The resulting function, called the cumulative amplitude function (CAF; Fig. 1), shows how CM amplitude grows as more of the cochlear partition becomes unmasked and available to respond. In 2012, Chertoff et al. were able to use the CAF to predict locations of OHC loss in gerbil ears damaged with high-level noise exposure or laser light ($r = 0.734$, $p < .01$).

In spite of its fairly good ability to assess OHC loss, there were two apparent problems with this approach. Literature suggests that a significant portion of a RW-recorded response would be comprised of AN activity (Patuzzi et al., 1989; Henry, 1995; He et al., 2012; Lichtenhan et al., 2014). In 2014, Chertoff et al. set out to quantify the contribution to the RW response. In order to accomplish this, they damaged the AN with ouabain, a sodium-potassium ion pump inhibitor, and compared the responses before and after damage and between ouabain-treated and control groups. They reported that up to 70% of the response to a 762 Hz tone at low-moderate intensities could be basal AN activity. The second problem with 762 Hz CAF, was the limitation inherent in its shape. As shown in Fig. 1 (solid line) the CAF has a sigmoidal shape and the apical-most portion of the function is relatively flat and close to the noise-floor, compressing any potential reduction in the curve due to damage in the apical region of the partition Fig. 1 (dotted line). The methods of Charaziak et al. (2017) ameliorated this problem by implementing a different method of recording location-specific CM. Charaziak and colleagues varied the stimulus frequency while using secondary suppressing tones slightly higher- or lower-than the stimulus frequencies in order to record the CM from specific locations.

Kamerer et al. (2016) acknowledged that the response recorded from the RW was likely the combination of responses from at least OHC and basal AN and thus referred to this response as the cochlear response (CR) to avoid confusion with the CM, which is purely an OHC response. The potential recorded from the RW will thus be referred to as the CR for the purpose of this paper. To further Chertoff's research into the use of the CR to assess the health of OHCs, Kamerer et al. attempted to avoid a neural contribution and record a pure CM by lowering the stimulus frequency to 45 & 85 Hz, which in the gerbil, was thought to be below the excitatory frequency range for the AN. The theory was that these frequencies would elicit motion in the BM and thus produce current flow through OHCs, but not elicit AN activity. Furthermore, Kamerer et al. hypothesized that lowering the stimulus frequency would result in a more linear CAF that could reveal potential damage to the apical regions of the cochlea (Fig. 1[b]), resolving the second issue with the 762 Hz CAF. Kamerer and colleagues repeated the ouabain experiments completed by Chertoff et al. (2014) using the 45 & 85 Hz stimuli. They did not find a

significant change in the CAF after application of ouabain and loss of the basal AN, and concluded that there was no significant basal AN contribution to the low-frequency CR. Thus, their recommendations were that low frequencies such as 45 & 85 Hz would be optimal stimuli to use in order to assess the health of OHCs. As to their second hypothesis, the 45 & 85 Hz CAFs maintained the sigmoidal shape and Kamerer et al. did not recommend that these stimulus frequencies could be used to assess the health of apical-most OHCs. While the CAF of the low-frequency CR did not change much, the response waveform itself proved to be unique to the higher-frequency CR, and formed the impetus for the current study.

The long-term goal of this research is to develop a new diagnostic tool for sensorineural hearing loss that has the capability of assessing the health of anatomic structures in the cochlea that are thought to be most prone to damage and dysfunction. Additionally, locating the spatial site of damage along the cochlear partition is necessary for accurate treatment and prevention of unintentionally damaging healthy sections of the partition. The aforementioned Chertoff and Kamerer studies focused on tailoring the CR to record responses from OHCs specifically. This study takes this idea one step further and asks: What are the cellular sources of the low-frequency CR? With the future goal of a potential tool that could assess the health of *multiple* structures of the cochlea, not just OHC.

Kamerer and colleagues showed that the low-frequency CR does not contain significant contributions from basal AN, so a response to a tonal stimulus that is entirely from OHCs should mimic the stimulus both spectrally and temporally. In the case of 85 and especially 45 Hz, however, the response waveform (Figs 3-6) is distorted almost beyond recognition of the stimulus waveform (Fig 2). These large deflections in the waveform were found to be consistent across animals, suggesting additional cellular contributors to the response, beyond that of OHC. Additionally, the morphology changed drastically and in a patterned manner as a function of the conditions of the masking paradigm, suggesting that the contributions of these structures to the response change in relation to each other as a function of masker cutoff frequency. The purpose of this study was to identify the cellular contributions to the low-frequency CR, to determine its potential as a diagnostic tool that is objective in nature and could be used to assess

the health of multiple anatomical structures in the cochlea at discrete spatial locations along the cochlear partition. The study is divided into four experiments, and the results of each provide direction for the next.

First, it was necessary to confirm that the rich harmonic distortion in the low-frequency CR was produced in the cochlea. Experiment 1 tested the hypothesis that the distortions seen in the low-frequency CR were due to nonlinear mechanisms in the cochlea, and not simply a reflection of distortions present in the acoustic stimulus. Different cellular structures in the cochlea produce unique patterns of distortion depending on location and stimulus conditions, thus analysis of the harmonic components of the low-frequency CR may provide clues as to the cellular sources contributing to response. Experiment 2 looks at each distortion-product, or harmonic, as a function of masking condition to observe their relationships with the fundamental frequency response and each other. There are a number of suspected cellular contributors to the low-frequency CR; Physiologic experiments which eliminate a suspected source and observe the effects on the response are a definitive method to identify and quantify these cellular contributions. The third experiment eliminates the basal auditory nerve to quantify its contribution to the low-frequency CR. Another method of identifying these sources without the use of animal experiments is mathematical modeling: Experiment 4 uses two analytic approaches, independent component analysis and linear regression, to identify the sources and direct future physiologic experiments.

GENERAL MATERIALS & METHODS

Animal Preparation

Eighteen adolescent Mongolian gerbils (Meriones unguiculates) weighing between 40 and 70 g were used for this study. The gerbil model was chosen for its low-frequency hearing sensitivity, allowing for potential translatability of these methods to the human ear. Prior to surgery, animals were sedated with an intraperitoneal injection of pentobarbital (64mg/kg). The animals were laid in a supine position on a stereotactic platform with their front teeth hooked into a head-stabilizing mechanism, and wrapped in a heating blanket (Harvard, Hollistan, MA) to maintain a body temperature of 37° C. Sedation was maintained throughout the experiment with hourly supplemental intramuscular injections of one-third the initial dose and monitored by heart-rate and pulse-oxygen levels every 10–20 minutes (MouseOx, Starr Life Sciences, Oakmont, PA). Following a local injection of lidocaine, the right pinna was removed and a post-auricular incision made to expose the bulla. A pick was used to open a small hole in the bulla over the RW niche, just large enough to insert the electrode and a cotton wick to absorb any condensation and excess fluid. All experiments were acute, in that they were completed on a single animal within several hours followed by sacrifice with three times the initial dose of pentobarbital. All experimental procedures were approved by the University of Kansas Medical Center Institutional Animal Care and Use Committee.

Generation of Stimuli and Masking Noise

The stimuli used to elicit the CR in this study were 45 & 85 Hz tonebursts. The tonebursts were generated in BioSig (Tucker-Davis Technologies) to have a 70 ms duration and 15 ms rise/fall time and were presented at a rate of 5/s to a speaker located outside of a soundproofed chamber. The 12 cm-diameter speaker was coupled to a tube which entered the chamber and terminated at the animal's external auditory meatus. The tube was 1.9 m long to facilitate resonance of the 45 Hz stimulus and reduce acoustic distortion and electrical artifact. The masking noise paradigm was a series of broadband noise filtered with seventeen high-pass cutoff frequencies, generated by a real-time processor (TDT, RZ6). The masking noise was delivered simultaneous to the stimulus by a separate speaker placed

approximately 3 mm from the ear canal (TDT, MF1). The frequency response of the speaker was flattened from 0.1-60 kHz via an FIR filter. A ¼ inch Bruel and Kjaer microphone was suspended within several millimeters to the ear canal and the end the speaker tube. The frequency sensitivity of the B&K microphone was flat (+/- 2 dB) from the 0.02-100 kHz when positioned at a 90° azimuth relative to the sound source. The microphone output was routed to an oscilloscope (Tektronix, TDS 2014) to calibrate the peak sound pressure level of the toneburst stimuli.

CAPs were also recorded from these animals to check for hearing loss prior to recording the CR. CAPs were evoked with 1, 2, 4, 8, 16, and 24 kHz alternating-polarity tonebursts with 1 ms rise/fall times and an 8 ms plateau. Each frequency was presented 50-500 times at 21.1/second starting at 100 dB SPL and descending in 5 dB steps until threshold was reached. The CAP tonebursts were presented from the same speaker used to present the masking noise for the CR experiments (TDT, MF1) and sound pressure levels were calibrated using the B&K microphone.

Physiologic Recordings

The CR and CAPs were recorded via a ~1 mm diameter silver-wire ball electrode, placed with a micromanipulator device into the RW niche of the exposed bulla. A needle electrode placed in the neck muscle served as a ground. The CR was recorded to 45 & 85 Hz tonebursts presented at both 80 & 90 dB SPL, amplified 50x (Stanford), sampled at 200 kHz (TDT, RZ6), filtered from 0.03-30,000 Hz, amplified a second time (100x), and low-pass filtered from 3 kHz (Stewart). The CR was averaged over 50 recordings at each masking condition and each masking condition was recorded three times per animal, for an averaged CR of 150 recordings per stimulus frequency and level for every masking condition. The unmasked condition was recorded three times: once before the masking noise, once after the 9th cutoff frequency (to check for posterity), and once after the last cutoff frequency, for an average of 400 recordings per animal per stimulus condition.

Procedures

Once the animals were sedated and the bulla opened, CAPs were recorded to determine normal hearing sensitivity and the CR was recorded in healthy animals. The four stimulus conditions (45 & 85

Hz, at 80 & 90 dB SPL) were presented in random order. The masking noise conditions were recorded in order of lowest cutoff frequency to highest with three unmasked CRs recorded as well. All stimulus conditions and masking conditions were recorded from each of the eighteen animals.

EXPERIMENT 1: ACOUSTIC ANALYSIS

The striking distortion visible in the CR waveform, especially at 45 Hz, was originally a cause for concern with the stimulus. Delivering such low-frequency tones with high-fidelity required a relatively large speaker (9 cm diameter), which, in order to direct into the gerbil ear canal, had to be funneled into a 7 mm-diameter tube. We suspected this funneling would cause some amount of distortion in the acoustic stimulus, but not as much as was found in the response. Our hypothesis was that the amount of harmonic distortion found in the animal CR was greater than the expected response to the levels of distortion present in the acoustic stimulus, meaning the distortions were products of nonlinear mechanisms within the cochlea. To test this hypothesis, we implemented a strategy similar to that Dallos & Sweetman (1969) used to compare the spatial distribution of harmonics in the CM, in order to show that the generation of the harmonic components were not simply a response to acoustic distortion.

Experiment 1: Analysis

The 45 and 85 Hz acoustic stimuli were recorded at the highest levels used in this study (90 dB SPL) at the level of the ear canal –i.e. after passage through the funneling delivery apparatus –with a B&K microphone. The stimulus time-domain signal was converted to the frequency domain using the Fast-Fourier Transform (FFT) and the magnitude of the peaks of the first two harmonics, F2 and F3, were measured. Each of these harmonics found in the 45 & 85 Hz stimuli were used to create four individual toneburst stimuli (90, 135, 170, & 255 Hz) that were presented at the corresponding levels measured within the 90 dB SPL original acoustic stimuli (50, 47, 54, & 47 dB SPL, respectively). These stimuli were each in turn presented to an animal and the elicited CRs were recorded at all masking noise conditions in the same manner as explained in the general Materials and Methods. Using the same animal, the CR for 45 and 85 Hz at 80 dB SPL were also recorded.

Experiment 1: Results

It was necessary to determine whether the harmonic distortion seen in the low-frequency CR was a product of the nonlinear mechanisms within the cochlea, or simply a linear response to distortion in the acoustic system due to a delivery system susceptible to distortion-generation. In order to do this, we

recorded one animal's responses to tones with the same frequency as the harmonics found in the 45 & 85 Hz CR. The FFTs of the animal's CR to 45 and 85 Hz are shown in Fig. 7. The amplitude of the F2 and F3 harmonics present in the CR to 45 and 85 Hz, 80 dB SPL tones (arrows) were measured and compared to the amplitude of the response to the four tonal stimuli of the same frequency presented at the peak sound pressure levels within the 45 and 85 Hz acoustic stimuli. The amplitudes of all responses from this animal are plotted in Fig. 8 as a function of distance from the apex –derived from the cutoff frequencies of the filtered noise conditions. The functions show the growth of the harmonics as more of the cochlea becomes available to respond with increasing cutoff frequency. The blue and orange lines represent the harmonics of the 45 (top) and 85 Hz (bottom) CR, while the yellow and purple lines represent the response to those tones at the same frequencies at the levels they are present within the 45 and 85 Hz acoustic stimuli. It is clear that harmonic distortion found within the 80 dB SPL 45 and 85 Hz cochlear response is much greater than the responses to the tones of the same frequency, even when presented at levels found in the higher level (90 dB SPL) stimuli. This is evidence that these large harmonic distortions are indeed generated by nonlinear mechanisms within the cochlea.

Experiment 1: Discussion

The low-frequency CR to 45 & 85 Hz is unique to the higher-frequency CR because of the dramatic distortion seen in the waveform. At first, we were concerned that the stimulus delivery system was the culprit of said distortion that was just being reflected by OHCs in the CM, however, the distortions present in the CR seemed larger than the amount of distortion in the stimulus waveform. We hypothesized that, in fact, this distortion originated from nonlinear mechanisms in the cochlea. To resolve the question of origin, we recorded the 45 & 85 Hz cochlear response as well as responses to tones of the same frequencies as the harmonics found in the CR. The amplitudes of the harmonic distortion components were much greater than the response to tones of the same frequency, even when comparing harmonic components in the lowest stimulus level CR (80 dB SPL) to the same frequency tones at levels in the highest stimulus condition (90 dB SPL). The data argues that the harmonics found in the low-frequency CR are indeed distortion-products of the nonlinear mechanisms in the cochlea, and can be

supported by the literature surrounding nonlinearities in the hair cell transduction mechanism and the origins of the CM. (see *Appendix A*). The CR is dominated by the CM, an extracellular measurement of changes in hair cell receptor potentials; mainly OHC. At high stimulus levels like those used in the present study, the output of the cell is likely being elicited from the saturated or compressive portion of the cell's transducer function. This produces distortion in the response.

Outer hair cells typically have a symmetrical transducer function and produce odd higher-order harmonics (i.e. F3, F5, etc.). There were large amounts of F3 and even F5 in the CR to both 45 & 85 Hz tones, suggesting one source of the CR to these tones can be identified as OHC, which is to be expected and substantiated by the literature (Dallos & Durrant, 1972; Dallos, 1973; Dallos & Cheatham, 1976). In addition to odd harmonics, significant even-order harmonics were found in the low-frequency CR, especially at 45 Hz. Inner hair cell transducer functions are typically asymmetric which produce both even and odd harmonics (F2, F3, F4, F5, etc.). The CM is the reflection of current flow through hair cells, including IHCs, though the CM is highly attributed to OHCs and is not used to assess IHC physiology. However, the presence of significant even order harmonics suggests that IHC could also be a significant source of the low-frequency CR. This idea was intriguing, as the majority of studies on the IHC are done *in vitro* or via invasive *in vivo* measurements in animals, and there is no current clinical measure for assessing the health of IHCs, despite their role as the primary afferent sensory cell of hearing.

We concluded from this experiment that the low-frequency CR has large higher-order harmonic distortions that originate from biological processes within the cochlea.

EXPERIMENT 2: DISTORTION-PRODUCTS OF THE COCHLEAR RESPONSE

Due to different mechanisms of excitation, inner and outer hair cells produce different harmonics (see *Appendix A*). Additionally, the auditory nerve produces patterns of harmonics depending on their location of generation and health. Thus, it was important for us to analyze the spectral content of the low-frequency CR to obtain clues as to its cellular sources.

Experiment 2: Analysis

The CR time-domain signal was windowed with a Hanning window, then converted to the frequency-domain using the FFT (MATLAB). The magnitude of the peaks of the following spectral components were considered the amplitudes for each harmonic: the fundamental frequency (F0), either 45 or 85 Hz; ~3 Hz, the frequency of the stimulus envelope (Env), was included in the harmonic analyses (though a distortion-product rather than a true harmonic); second harmonic (F2), and third harmonic (F3). The amplitudes of the fundamental and each harmonic were plotted separately as a function of distance from the apex, converted from masking cutoff frequency, in the same manner as the CAF described by Chertoff et al. (2012, 2014), with the exception that we did not normalize to the unmasked amplitude. This resulted in separate growth functions for each harmonic (F0, Env, F2, & F3) for each fundamental frequency (45 & 85 Hz), sound pressure level (80 & 90 dB SPL), and animal. In order to make more accurate comparisons across animals, the ratio of each harmonic to the fundamental was calculated within each animal (n = 16), which resulted in growth curves of harmonic ratios (Env/F0, F2/F0, & F3/F0) for each stimulus condition. A linear mixed-model was regressed to the data, and because of the sigmoidal shape of the growth functions, a third order polynomial was an appropriate fit. In order to determine the significance of harmonic ratio to the shape of the growth curve, a linear mixed model was fit to the data:

$$y_i = \alpha + \beta_1 d + \beta_2 d^2 + \beta_3 d^3 + \beta_4(harmonic) + \beta_5(d \times harmonic) + \beta_6(d^2 \times harmonic) + \beta_7(d^3 \times harmonic) + \varepsilon \quad (1)$$

where y is the harmonic ratio growth curve, d is the distance from the apex (in mm) along the cochlear partition, β is the regression coefficient, and *harmonic* is the categorical variable for the three harmonic ratios. The variables $\beta_5 - \beta_7$ represent the interaction between the distance variables and harmonics.

Backward elimination was applied to the model to reach a final model for each stimulus condition (45 & 85 Hz, at both 80 & 90 dB SPL) that held only terms that were significant ($p < 0.05$). Dummy variables were then made for each harmonic ratio, in order to develop a model capable of comparing the growth curves of each ratio. The full model is shown below.

$$y_i = \alpha + \beta_1 d + \beta_2 d^2 + \beta_3 d^3 + \beta_4\left(\frac{F2}{F0}\right) + \beta_5\left(\frac{F3}{F0}\right) + \beta_6\left(d \times \frac{F2}{F0}\right) + \beta_7\left(d \times \frac{F3}{F0}\right) + \beta_8\left(d^2 \times \frac{F2}{F0}\right) + \beta_9\left(d^2 \times \frac{F3}{F0}\right) + \beta_{10}\left(d^3 \times \frac{F2}{F0}\right) + \beta_{11}\left(d^3 \times \frac{F3}{F0}\right) + \varepsilon \quad (2)$$

where y is the growth curve, d is the distance from the apex (in mm) along the cochlear partition, β is the regression coefficient, and F2/F0 & F3/F0 are represented as 1 to solve the equation for that ratio, and set as 0 otherwise, so that the equation for Env/F0 was represented when both F2/F0 & F3/F0 were equal to 0. Backward elimination was applied to the full model to reach a final model for each stimulus condition (45 & 85 Hz, at both 80 & 90 dB SPL) that held only terms that were significant ($p < 0.05$). To make comparisons between F2/F0 and F3/F0, the reference dummy was changed from Env/F0 to F2/F0, meaning the equation was changed such that if Env/F0 & F3/F0 were set to zero, the equation would solve for the regression line for F2/F0. SPSS software was used to fit the model (StataCorp LP, College Station, TX).

Experiment 2: Results

Inner and outer hair cells and the auditory nerve will produce patterns of harmonic distortion due to stimulus condition and differences in their excitatory drives. Therefore, it was of interest to us to observe the patterns of the different harmonics seen in the whole CR. One way to examine these harmonics was to compare differences in their growth patterns as a function of distance along the cochlear partition. The theory of the masking paradigm is that the response recorded at a given masking cutoff frequency is primarily generated by cells whose characteristic frequency is at or below the cutoff frequency of the masking noise. If this is assumed, then the harmonics can be plotted as a function of distance from the apex corresponding to the masker cutoff frequency. In Fig 9, we have done just that for each harmonic ratio (Env/F0, F2/F0, & F3/F0) at each stimulus frequency and level. First, the overall

significance of harmonics on the growth curves was determined. Table. 1 reports the coefficients and 95% confidence intervals of the harmonic variable and each of the distance variables. The p-values reported under each distance variable indicate significant two-way interactions between harmonic ratio and the linear distance; distance-squared, which explains a quadratic shape to the growth curve; and distance-cubed, which explains the sigmoidal shape to growth curve, for all stimulus frequencies and levels. Harmonic had a main effect on the growth curve at the higher stimulus level (90 dB SPL) at both frequencies.

Then the second regression model (Eq. 2) was fit to make comparisons across harmonics. The calculated individual regression lines for each harmonic ratio are plotted over the data in Fig. 9. The regression lines of the individual harmonic ratios were compared with each other and the t- and p-values of those comparisons are shown in Table 2. Paired comparisons having any significant two-way interactions, suggest differences in the shape of the growth curve. The distance variables of the Env/F0 harmonic ratio were significantly different from those of the F2/F0 and F3/F0, meaning the shape of the growth curve for Env/F0 differed from that of F2/F0 and F3/F0. The F2/F0 growth curve was different from F3/F0 only at the 85 Hz, 90 dB SPL stimulus condition. At all other stimulus frequencies and levels, there was no significant interaction between F3/F0 and the distance variables in reference to F2/F0, suggesting that the shape of these growth curves are not statistically different. Although there was no significant interaction terms, the 45 Hz, 90 dB SPL and 85 Hz, 80 dB SPL had significant main effects of F3/F0 with reference to F2/F0, meaning that there was a general amplitude difference between the two harmonic ratios (45 Hz, 90 dB SPL [$t = 10.628$, $p = 0.000$]; 85 Hz, 80 dB SPL [$t = 4.745$, $p = 0.000$]). At 45 Hz, 80 dB SPL, there was no statistical difference between the F2/F0 and F3/F0 growth curves (main effect of F3/F0: $t = -0.424$, $p = 0.672$).

Experiment 2: Discussion

The masking paradigm developed by Chertoff and colleagues (2012, 2014) was used to obtain the CR from specific regions of the cochlear partition. We noticed that the low-frequency CR morphology was changing as a function of high-pass masking noise cutoff frequency. Upon inspection of the spectral

components of the waveforms in masking noise, we hypothesized that the amplitude of different harmonic components changed as a function of masker cutoff frequency and did not share the same growth patterns. A mixed regression analysis of the growth functions of each harmonic component and post hoc comparisons between the components were performed for each stimulus frequency and level. The regression analysis revealed that the growth functions of the harmonics with reference to F0 were dependent on masker cutoff frequency, and each component was modeled with a third-order, sigmoidal function. Since each harmonic component was a ratio of its amplitude with reference to the fundamental frequency, this result means that the harmonic components change relative to the F0 of the CR as more of the cochlear partition becomes available to respond.

The Env/F0 ratio had a pronounced quadratic shape to the growth curve in every stimulus condition; the amount of the Env component increased as cutoff frequency increased, until approximately the 3700 Hz cutoff frequency, corresponding to 5 mm from the apex, where it peaked at about two times the amplitude of the F0. After this masking condition, as more of the cochlea became available to respond and the F0 increased, the Env component decreased. The source of the Env component is arguably neural or hair cell. The envelope-following response (EFR) is a steady-state evoked potential entrained to the envelope of the stimulus at frequencies up to ~1 kHz (Dolphin & Mountain, 1991). The lower the modulation frequency, the larger the response. A toneburst, like the stimuli used in this study could be thought of as amplitude modulated tones, in this case, with a modulation frequency of ~3 Hz, which may explain why there is such a large Env component in the low-frequency CR. Interestingly, the generators of the EFR at such low modulation frequencies are thought to be neurons in the brainstem, however, these conclusions are based on scalp-electrode recordings (Dolphin & Mountain, 1992; Kuwada et al., 2002; Shaheen et al., 2015). The auditory nerve and the hair cells also have an EFR and would be the most likely contributors to the response recorded from a RW electrode. Researchers who study the EFR to diagnose auditory neuropathy generally prefer scalp recordings *because* the EFR from hair cells and the peripheral auditory nerve may not be distinguishable (Shaheen et al., 2015). The Env component was largest for the lower stimulus level conditions (80 dB SPL) and was significantly different from the

higher-order harmonics at all stimulus conditions, possibly suggesting a different mechanism of origin for this distortion-product than that of the higher-order harmonics, or a combination of multiple contributors.

The similarity of second and third harmonics with reference to the fundamental frequency were dependent on stimulus frequency and level. At some stimulus conditions, these higher-order harmonics shared a large quadratic shape in their growth curves like the Env, meaning they had a steeper growth as a function of masker cutoff than did F0. At other stimulus conditions, these harmonics had relatively flat growth functions, meaning they grew at a more comparable rate to the F0. It is difficult to conclude that the source of F2 and F3 are different in terms of their cellular origin, given the results of this experiment. Conversely, it is just as difficult to conclude that they originate from the same source. While they do share similar features in terms of their growth patterns in the presence of masking noise, they are not identical and are statistically different in some stimulus conditions.

Regardless of stimulus condition, the mid-cutoff frequency masking conditions produced the largest harmonic ratios. This is not due to an increase in harmonic amplitude, but rather a reduction in F0 amplitude. The CAF, or growth function of F0, at 45 Hz is plotted in Fig. 4 of Kamerer et al. (2016). At the lowest cutoff frequencies, corresponding to the apical-most regions of the cochlear partition, the amplitude of F0 increases until 4 mm from the apex where amplitude begins to decrease and reaches a local low at 5 mm before it begins to increase again at 6 mm from the apex. This corresponds exactly with the peak of the harmonic ratio suggesting that the magnitude of the harmonics actually remain fairly constant while the fundamental frequency reduces in amplitude. Reduction of the CM is known to occur when the electrode is proximal to the location of maximal BM displacement, as this location produces the greatest phase rotations, and the responses from out-of-phase components are cancelled. However, the electrode site for these experiments are the RW and the stimulus frequency low. This should cause maximal BM displacement in the apical-most regions of the partition, quite distal to the electrode in the base where BM movement is, in theory, moving in-phase. Therefore, this explanation does not make sense. The most likely explanation for the reduction in F0 amplitude is that there are multiple cell populations contributing to F0 in this region of the partition corresponding to the dip in the CAF, that are

out-of-phase with each other, resulting in cancellations of their responses at F0. Whether these out-of-phase populations are OHC and AN, IHC, or different populations of OHCs could not be determined at this point.

EXPERIMENT 3: PHYSIOLOGIC APPROACH TO IDENTIFYING SOURCES

The harmonics found in the low-frequency CR are of great interest as they provided some evidence for multiple cellular origins. The logical next steps toward identifying the cellular contributors are physiologic experiments which remove a suspected source and observe changes in the response. One of the suspected sources was the auditory nerve, which could be contributing to the Env if indeed the contributor of the EFR in neural. The experiments of Chertoff et al. (2014, 2015) used ouabain to damage the basal auditory nerve and reported the effects on the fundamental frequency component of the CR; hence we were familiar with the procedures involving ouabain and the physiologic extent of damage that was to be expected as well as the minor effects on basal OHCs. Therefore, we began exploring the sources of the distortion with the basal auditory nerve.

In this experiment, the contribution of the basal AN to the harmonics of the low-frequency CR was determined by impairing the basal AN and observing the effects on each harmonic. To achieve this, a group of animals received a treatment of ouabain. Ouabain is a Na^+-K^+-ATPase pump inhibitor which was chosen based on data from several studies indicating its effectiveness in knocking out the AN (Schmiedt et al., 2001; Lang et al., 2005; McLean et al., 2009; Yuan et al., 2014; Chertoff et al., 2015; Kamerer et al., 2016). Sodium-potassium ion pumps are found in cells throughout the cochlea, including the AN, hair cells, and strial cells, however, ouabain can be used to selectively damage the AN because of its affinity for the α3-receptor subunit isoform (O'Brien et al., 1994; Pierre et al., 2008) expressed only by AN fibers (McLean et al., 2009) and its low affinity for the α1-receptor isoform found in the epithelial cells of the cochlea, including hair cells (McGuirt & Schulte, 1994). Several studies have also found that the sensitivity to ouabain is dependent on nerve type, with type I neurons having a higher sensitivity than type II (Lang et al., 2005), and low spontaneous-rate fibers having a higher sensitivity than high spontaneous-rate fibers (Bourien et al., 2014). While effective at damaging auditory nerve fibers, the dosage, treatment method, and application time of ouabain are important parameters when trying to avoid damage to hair cells. Bourien et al. (2014) measured the physiologic effects of ouabain on OHCs by recording distortion-product OAEs (DPOAEs) after an acute application of 100 µM ouabain on the gerbil

RW for 30 minutes. They found that ouabain did not affect the DPOAE amplitudes in the f_2 range of .5 – 20 kHz at 60 and 55 dB SPL signal levels for f_2 and f_1, respectively. Schmiedt et al. (2001) showed similar results for 1 mM ouabain, but found that 10 mM significantly reduced DPOAE amplitude after three hours. On the contrary, Chertoff et al., (2014) and Kamerer et al., (2016) did not find significant differences between 10 and 1 mM ouabain (respectively) on the DPOAEs of gerbils after acute applications on the RW for 30 minutes. They reported no effects on DPOAEs to 1-8 kHz tones and ~5 dB reduction in DPOAE amplitude at 16 & 24 kHz. The chosen dosage of ouabain for this study, therefore, was 1 mM.

Experiment 3: Procedures

Animals were designated to either a control or experimental group. After the original CAP and CR recordings, animals then received either the control or experimental treatment. Animals in the control group (n = 9) received a solution of artificial perilymph (120 mM NaCl, 3.5 mM KCl, 1.5 mM $CaCl_2$, 5.5 mM glucose, and 20 mM HEPES, with an adjusted pH of 7.5 via NaOH). Animals in the experimental group (n = 9) received 1 mM ouabain dissolved in artificial perilymph. Approximately 1–2 µL of either solution was dripped into the RW niche. After 30 minutes, the niche was dried with a cotton wick, and after an additional 30 minutes, the CAPs and CR were recorded again.

Experiment 3: Analysis

Kamerer et a. (2016) used a regression approach to determine the effect of 1 mM ouabain on the CAF –effectively the F0 growth curve –for the 45 & 85 Hz CR, and found that the loss of the AN did not significantly affect the F0 growth curve. With this in mind we decided to also examine any possible effects of ouabain on the harmonics, by comparing the harmonic components between ouabain- and AP-treated animals (n = 9 per group). A regression approach like that used to analyze the harmonic ratios in Experiment 2 was performed in this experiment to compare the effect of ouabain across the harmonic components. Instead of ratios, the difference between before and after treatment was calculated for each spectral component (Env, F0, F2 & F3) within each animal. The difference values as a function of distance did not show as much of a shape like that of the harmonic ratios, which required a third-order

polynomial equation to fit the growth curve, therefore, Eq. 1 was modified to a linear function (d^2 and d^3 were removed) and a treatment variable was added.

In addition to making comparisons across harmonics at each stimulus condition individually, the effect of ouabain on each harmonic across stimulus frequency and level was analyzed to determine if the effect of ouabain on a given harmonic was dependent stimulus condition. Equation 3 (below) shows the model for this regression analysis.

$$y_i = \alpha + \beta_1 d + \beta_2(TX) + \beta_3(freq) + \beta_4(level) + \beta_5(d \times TX) + \beta_6(TX \times freq) + \beta_7(TX \times level) + \beta_8(d \times TX \times freq) + \beta_9(d \times TX \times level) + \varepsilon \quad (3)$$

where y_i is the difference between before and after treatment of one harmonic (Env, F0, F2, or F3), *TX* is the treatment variable (ouabain or artificial perilymph), *freq* is the stimulus frequency variable (45 or 85 Hz), and *level* is the stimulus level (80 or 90 dB SPL).

Experiment 3: Results

In this experiment, we quantified the effects of ouabain on harmonic components to determine if basal auditory nerve was a source of the low-frequency cochlear response. To accomplish this, the difference in amplitude at each distance (derived from the cutoff frequency of the masker) before and after treatment was calculated for each animal. Figure 10 depicts the mean and 95% confidence intervals (n=9) for the difference before and after treatment (such that a negative value denotes a decrease in amplitude after treatment and vice versa), for each harmonic component (rows) and each stimulus condition (columns). Ouabain-treated animals are shown in red and artificial perilymph-treated animals in blue. A treatment variable was added to Eq. 1 from Experiment 2 and the data regressed separately for each stimulus condition with only a linear distance variable. In the 45 Hz, 80 dB SPL condition, there was no significant effect of ouabain on the harmonic components, including the fundamental (F0). At all other stimulus conditions, however, there were significant three-way interactions between treatment, harmonic, and distance ($p < .05$), meaning that at least one harmonic in each stimulus condition was affected by ouabain and this was dependent on masking condition/distance. Since there seemed to be a pattern of stimulus condition on the effect of ouabain, each harmonic component as a function of distance was

regressed with additional frequency and level variables (Eq. 3). The Env, F0, & F2 all showed significant three-way interactions between distance, treatment, and level, meaning that the effect of ouabain on these harmonics was dependent on stimulus level and distance from the apex ($p < .01$). In the higher-level stimulus conditions (90 dB SPL), the response from the basal regions of the cochlear partition were most affected by treatment and this was dependent on harmonic (Fig. 10, columns 2 & 4). The F0 and F3 show similar growth patterns: animals treated with artificial perilymph showed a slight decrease in amplitude and interestingly, those treated with ouabain showed a slight increase in amplitude (though changes in F3 were not statistically significant ($p >.05$). Conversely, the Env and F2 were positively affected by perilymph and negatively affected by ouabain. The data show that while statistically, ouabain seemed to affect some harmonics at some stimulus frequencies and levels, the overall impression of the data was that ouabain did not devastate any harmonic component.

Experiment 3: Discussion

The purpose of this experiment was to quantify any contribution of the AN to the distortion-products found within the low-frequency CR. The most definitive way to identify this contribution is to eliminate the auditory nerve and observe the effects on the response. In theory, the application of ouabain in the RW niche will eliminate the AN response, however, we learned from Chertoff et al. (2014) and Kamerer et al. (2016) that an acute application of ouabain on the RW for 30 minutes will not entirely eliminate AN responses below 8 kHz. Therefore, the conclusions we can draw from this experiment are limited to the contribution of the basal AN. Statistically there were several harmonics at some stimulus conditions that were affected by ouabain, however, no one harmonic component was devastated by the loss of the basal auditory nerve, eliminating the possibility that the basal AN is a major source of the low-frequency CR, especially when compared to the amount of basal AN contribution found in higher-frequency responses (Chertoff et al., 2014). It is interesting to note that the F0, and to a lesser extent the F3, experienced an increase in response amplitude after application of ouabain. Kamerer et al. (2016) speculated that the AN may have a small, but out-of-phase contribution to the CR that when removed, could lead to an increase in the response because it no longer subtracts from the OHC contribution.

The Env and F2 harmonic ratios experienced the opposite effect, with a slight decrease in amplitude after the application of ouabain which was dependent on stimulus condition. This may indicate a shared cellular source and one that is different from the source producing the F0 and F3 components. The negative affect of ouabain on the Env and F2 suggests that either the basal AN contributes a small amount to these components or that ouabain also affected whatever cellular source produced these components in addition to the AN. We know that ouabain has an affinity for the receptors found in the auditory nerve but many other cells in the cochlea have the receptor as well, including hair cells and strial cells. While the receptors in the epithelial cells of the cochlea are a different isoform than those found in AN fibers, they can still be affected by ouabain. It is also possible that the EFR contributing to the Env component seen in the CR had some peripheral AN origins. While it is difficult to definitively conclude the reason for these changes, none of the harmonics were drastically affected by the loss of the basal AN, suggesting this population of cells does not significantly contribute to the harmonic distortion seen in the low-frequency CR, nor the F0.

EXPERIMENT 4: ANALYTIC APPROACH TO IDENTIFYING SOURCES

The results of Experiments 1 & 2 suggest that the low-frequency CR is a complex waveform comprised of a mixture of response signals from more than one source. Experiment 3 provided evidence that the basal auditory nerve is not one of these sources. Physiologic experiments like Experiment 3 are expensive and require animals, but are truly the most definitive way to identify the cellular sources. This line of physiologic research is not complete with just this one experiment; however, instead of embarking on a wild goose chase of eliminating all the cells in the cochlea one-by-one in order to identify the contributions to the low-frequency CR, we employed a more strategic method. In this experiment, we used mathematical modeling to provide a compass for further physiologic experimentation by illuminating the most likely cellular sources. In Experiment 4 we implemented two statistical techniques to generate a mathematical model of the low-frequency CR and the cellular sources making significant contributions. We chose one stimulus condition to explore the effectiveness of these analyses; the 45 Hz, 80 dB SPL condition resulted in the largest amounts of harmonic distortion and had the largest change in the harmonic ratios as a function of distance, therefore we deemed this condition the best candidate for modeling.

Independent Component Analysis

First, we attempted a novel statistical technique to separate the CR waveform into its discrete source components. Blind source separation is a family of statistical methods capable of isolating discrete source signals from a mixture of signals, such as the low-frequency CR. Independent component analysis (ICA) is one method of blind source separation that works on the assumption that if different signals originate from different physical processes then the signals they produce are statistically independent. If these signals are combined in a linear manner they generate a mixture. This assumption can be reversed to take a mixture and find the statistically independent source signals that comprise it (see *Appendix B* for more information on the problem of blind source separation and the theory of ICA). Independent component analysis has never been used on auditory data like that of the CR, but we considered the low-frequency CR to be a mixture of independent source signals, so it seemed like an appropriate analysis for

the data. There are numerous kinds of ICA that use a variety of signal features and methods of estimating source signals. In this experiment, we used FastICA –a highly-efficient fixed-point iteration scheme that maximizes the nongaussianity of independent components within the mixtures (Hyvärinen & Oja, 1997; Hyvärinen, 1999) –in the form of a MATLAB algorithm.

The input for ICA is a time-domain mixture. We used the mean 45 Hz, 80 dB SPL CR waveform shown as the black line in Fig. 11 (top). Another important assumption of canonical ICA is that the number of input mixtures must equal the number of estimated sources, and these mixtures must be recorded simultaneously but separated spatially. We wanted ICA to estimate at least two independent sources, however, we only had one electrode location (the RW), and consequently, only one mixture. We attempted to overcome this deficit by using two consecutive recordings in the same animal, however, we discovered that the reasoning behind the assumption that the recordings must be spatially separated as to generate two mixtures that contain slight differences. Spatially separated mixtures will have both spectral and temporal differences as the source signal amplitudes and phases are weighted by their distance to recording mechanism, in our case: the electrode. Our solution was to change the stimulus phase between the two consecutive recordings, which would provide at least temporal differences between the two mixtures. The mean and all responses to the flipped stimulus are shown in Fig. 11 (bottom).

Before being analyzed by ICA, the mixtures were preprocessed. The data were centered and whitened using a principle component analysis algorithm (MATLAB). The basic parameters of the algorithm were selected, that is, the approach: either symmetric, which estimates components simultaneously, or deflation, which estimates the components one-by-one; and the number of estimated sources: two. Additional parameters can be entered into the algorithm which aid the analysis in searching for the source components by giving it some known information about the sources. In this case, the nonlinearity 'g' used in the algorithm was set to:

$$g(u) = \tanh(a \times u)\,; a = 1 \tag{4}$$

to estimate the cumulative distribution function of a sinusoidal signal. More information on preprocessing and algorithm parameters can be found in *Appendix B.*

ICA Simulation

Before analyzing the real CR data, an overly-simple simulation of a CR was created to test if ICA could perform well on CR-type data. In order to recreate CR-like data, we simulated two possible cells in the cochlea that could produce the harmonics we saw in Experiment 1: inner and outer hair cells. Real data from Russell, Cody, & Richardson (1986) were taken to simulate hair cell responses. In their paper, intracellular I/O functions were created from acoustically-stimulated inner and outer hair cells. These I/O functions, or transducer curves (or functions) as we refer to them, for an inner and outer hair cell are shown at the top of Fig. 12. Transducer curves plot the cell's output (Volts) when stimulated with a given pressure (Pascals). The 45 Hz, 80 dB SPL stimulus that was recorded in Experiment 1 was then used to "drive" the simulated OHC transducer curve and the differentiated stimulus was used to "drive" the IHC transducer curve, resulting in simulated response waveforms from an IHC and OHC. A more in-depth explanation of the Russell, Cody, and Richardson data that was used and the methods of simulating hair cell responses will be provided by the linear regression model, where this simplified simulation is elaborated and refined. The IHC and OHC response waveforms became our 'known' sources for this test of ICA. In order to create mixtures to be analyzed by ICA, the two simulated hair cell responses, or sources, were arbitrarily weighted and mixed into two mixtures containing different amounts of each source. Mixture 1 contained three times the amount of OHC as IHC, and Mixture 2 contained equal amounts of both IHC and OHC (Fig. 12). These mixtures were the input for ICA and using the parameters specified in the *Analysis* (both the symmetrical and deflation approach worked), ICA was asked to recover two sources.

The estimated source waveforms are shown in Fig. 12 and a quick comparison of the estimated sources to the known sources proved that ICA competently recovered the simulated data. Upon closer inspection, a keen eye might note that the estimated OHC waveform is flipped 180°. This is due to ICA's inherent scaling ambiguity which includes a phase term, and therefore was ignored. The FFT of Mixture 1 and the two estimated source waveforms are plotted at the bottom of Fig. 12. It is of interest to note that the recovered source that estimated the simulated OHC had spectral energy at 0 Hz (a DC shift), 45 Hz

(F0), 135 Hz (F3), and 225 Hz (F5). The recovered source that estimated the simulated IHC had a larger DC amplitude, and spectral energy at 45 Hz (F0), 90 Hz (F2), as well as smaller amounts of 135 Hz (F3) and 225 Hz (F5). Again, a reader with an eagle eye might also see an energy spike at 180 Hz (F4), however, this is also the third harmonic of 60 Hz, where noisy electrical artifact resides, and therefore will be ignored in the discussion.

ICA Results

Once it was established that ICA could work on CR-like data, we set forth analyzing the real 45 Hz, 80 dB SPL data. Using the two out-of-phase mixtures in Fig. 11, we asked ICA to recover two sources. These estimated sources are shown in Fig. 13 in both the time-domain (top) and the frequency-domain (bottom). Morphologically, 'Recovered Source 1' is similar to the mean 45 Hz, 80 dB SPL CR waveform, and spectrally is also very similar in that the most energy resides at 45 Hz (F0), 135 Hz (F3), and 225 Hz (F5) with the exception of no DC shift. The 'Recovered Source 2' waveform clearly shows a low-frequency bias, or Env as we refer to it, but the waveform looks noisy. Interestingly, the FFT reveals there is significant energy at ~3 Hz (Env) but also at 90 Hz (F2) and 180 Hz (F4). So, while the recovered source waveforms morphologically do not seem to match our simulated hair cells, spectrally, they are very similar. Recovered Source 1 spectrally lines up with the simulated OHC and Recovered Source 2 is spectrally similar to the simulated IHC, with two exceptions: it is missing the 45 Hz (F0) component and the odd harmonics (F3 & F5).

Linear Regression Model

Independent component analysis was an approach to dismantle the low-frequency CR into its source components. In the second part of this experiment, we reversed this approach and built a model of the low-frequency CR using source components we created from real data found in the literature. A linear regression model was created in order to determine which covariates were significant to successfully model the real low-frequency CR data. The final model:

$$CR(t) = \alpha + \beta_1 OHCr + \beta_2 IHCr + \beta_3 ENV + \beta_4 AP + \varepsilon \quad (5)$$

where CR is the real mean cochlear response amplitude (V), $OHCr$ and $IHCr$ are the simulated outer and inner hair cell response amplitudes (V), ENV is the stimulus envelope amplitude (V), and AP is the simulated neural phase-locked action potential amplitude (V), at any given moment in time.

The input to the CR model was the 45 Hz stimulus, recorded at the level of the ear canal at 80 dB SPL (Fig. 2, top left panel). The stimulus waveform (*S*) was manipulated in order to provide the appropriate stimuli for both inner and outer hair cells. Thus, the waveform was differentiated once to simulate the BM response to stapes footplate velocity (S'), which is the input for OHCs. Because there is a small amount of distortion in the stimulus itself, the differentiated stimulus was smoothed over 200 pts to eliminate falsely-exaggerated distortion, and rescaled to .2 Pa (or 80 dB SPL). As the input to IHC is acceleration of the stapes footplate (or velocity of BM displacement) the stimulus was differentiated twice, smoothed, and rescaled to model the input for IHCs (S'').

***OHCr & IHCr* –Hair Cell Responses**

The model hair cell responses (*OHCr* & *IHCr*) were generated by driving the respective stimulus inputs through the transducer curves of each hair cell type, represented by the following set of equations:

$$\begin{aligned} OHCr &= f_{TrOHC}(S') \\ IHCr &= f_{TrIHC}(S'') \end{aligned} \tag{6}$$

where *S* is the stimulus and f_{Tr} are the transducer functions for OHC and IHC fit with a first-order Boltzmann function:

$$f_{Tr} = \frac{c}{1 + e\left(\frac{S - a}{b} + d\right)} \tag{7}$$

where *S* is the derived stimulus for either the OHC or IHC, and the coefficients *a-d* were determined by linear regression (MATLAB Curve Fitting). The data that was regressed to find f_{Tr} were taken from real data published by two groups of researchers: Russell, Cody, & Richardson and Dallos & Cheatham in guinea pigs. These scientists recorded the intracellular transducer functions of both IHC and OHC. Three IHC transducer functions were recreated from recordings from Dallos & Cheatham (1989) from an apical IHC with a characteristic frequency of 700 Hz, elicited with low-frequency tones; Dallos (1986) from a

third turn IHC, elicited with an 800 Hz tone; and Russell, Cody, and Richardson (1986) from a first turn IHC, elicited with a 600 Hz stimulus. Three OHC transducer curves were recreated with data from Dallos (1986) and Russell, Cody, and Richardson (1986) from OHCs proximal to their respective IHCs described above; and Cody & Russell (1987) from a first turn OHC elicited with a 600 Hz tone. Data were taken from figures of these publications and first-order Boltzmann functions from Eq. 7 were fit to each data set. The Boltzmann functions fit all six data sets well ($R^2 > .99$), and the solved equations were used as the transducer functions (f_{Tr}) to generate the simulated hair cell responses. The six transducer curves (three for each hair cell) are shown in Fig. 14(a-c & g-i). The notable differences in the data between the groups are 1) the locations from which the data was taken, i.e. basal v. apical hair cells; 2) the maximum amplitude of the hair cell outputs; 3) the slope of the linear portion of the function, which result in varying sharpness of compression phases; and 4) the symmetry of the functions, especially in the OHCs.

In order to generate the simulated hair cell time-waveform responses, we solved Eq. 6 using the derived stimuli (S' & S'') for each hair cell. The results are shown in Fig. 14 (d-f & j-l). Here one can see the relationship between the symmetry of the transducer functions and the DC shift in the resulting response waveform. The greater the asymmetry between depolarization and hyperpolarization of the hair cell, the greater the DC shift, e.g. the positive peaks have a greater absolute amplitude than the negative peaks, exemplified by the Dallos (1986) *OHCr* (Fig. 14[d]). The other two OHC transducer functions that are more symmetric yield response waveforms whose positive and negative peaks are more or less equal in amplitude and do not show morphological differences. The slope of the transducer function will determine the sharpness/steepness of the compressive saturating portion of the curve; a steeper function will saturate more sharply whereas a shallow slope may have a shallower compressive portion. The differences in the positive peaks of the inner hair cell response waveforms in Fig. 14(j-l) are a great example how a hair cell response will change morphologically (due to distortion) depending on where the stimulus level lies on the transducer function. In Fig. 14(l) the positive peaks of the .2 Pa (80 dB SPL) stimulus elicits a response from the linear portion of the transducer curve. This results in a response

waveform that does not show any flattening, or clipping, of the positive peaks, compared to the positive peaks of Fig. 14(j) which show a significant amount of peak clipping due to the stimulus eliciting a response from the saturated portion of the transducer curve.

***ENV* –Stimulus Envelope Response**

In the mid-cutoff frequency masking conditions, especially in the 45 Hz, 80 dB SPL stimulus condition, there is a large low-frequency bias to the CR waveform, calculated to be ~3 Hz, which shifts the CR in a positive direction. This EFR is thought to be either neural or hair cell, however, the results of Experiment 3 led us to believe this component may be the summating potential (SP) of one of the hair cells. The EFR for the regression model was simulated from the differentiated stimulus (i.e. BM response) and derived using the real values of the Hilbert transform (MATLAB). The final covariate, *ENV*, is shown in Fig. 16(c).

***AP* –Phase-locked Auditory Nerve Response**

One way that the low-frequency CR is unique to the higher-frequency (>150 Hz) CR, are the visually apparent action potentials that ride the waveform at consistent phase-locations. These APs were isolated by high-pass filtering the CR waveform at 300 Hz (MATLAB). Figure 15 is an example of this filtering in one animal at 45 Hz 80 dB SPL, where each row is the filtered CR of a noise condition: the marked rows are the unmasked condition, and the other rows from top to bottom go from the lowest cutoff frequency to the highest cutoff masking noise. The first five masking conditions (lowest cutoff frequencies) are too noisy to visually identify any APs, but beginning at the sixth, the APs are fairly clear. The AP wavelets were visually identified and agreed upon by three independent observers, using the following characteristics as guidelines: 1) the wavelet is *biphasic* as characterized by recording with extracellular electrodes, 2) the wavelet has a frequency of ~1 kHz, which is difficult to determine visually, and therefore we decided on a *steep slope* in the positive and negative directions as our visual determinant of this characteristic, 3) the wavelet is *periodic* in that it occurs regularly along the CR wave cycle (e.g. every 1 ms) and 4) is *maintained* across masking condition, i.e. once above threshold, remains in the same time location without jitter.

Although most animals had many different AP wavelets, there were two regular AP wavelets that were identified across all animals and stimulus conditions: One wavelet occurred at a 90° phase lag (AP+) and one at a 90° phase lead (AP-) on the CR waveform, i.e. on the positive and negative slopes of the waveform at ~.707 peak amplitude. These APs created a neural spike train, which heretofore we consider a 'pulse-train', of two times the stimulus frequency (i.e. 90 Hz pulse train for 45 Hz stimulus). This pulse-train was then simulated to make the *AP* covariate for the model. The *AP* pulse-train was created in MATLAB by filtering a 1 kHz sine wave to make a wavelet, repeating that wavelet at a frequency of 90 Hz, beginning with a time delay such that the first pulse initiated at a 90° phase lag to the CR waveform. The finalized covariate, *AP* for a 45 Hz signal, is shown in Fig. 16(d).

Linear Regression Results

The linear regression model in Eq. 5 was implemented to estimate the coefficients of the model covariates and determine their significance to the model, with the intent of quantifying the contribution of each cellular structure's response to the CR to a 45 Hz, 80 dB SPL tone burst. There were three simulated hair cell responses each for inner and outer hair cells. The best pair to use in the model was determined via R^2, or model fit to the real data. The R^2 values of model fit using each possible pair of *IHCr* and *OHCr* are displayed in Table. 3. Two R^2 values are reported for each pair: one for the CR in the unmasked noise condition (top) and the second for the masking condition with the cutoff frequency of 3700 Hz (bottom), which is the condition that results in the largest morphological deviation from the unmasked waveform. The best fit of the model to the real CR for both the unmasked and masked noise conditions were using the OHC data from Dallos (1986) and the IHC data from Russell, Cody, and Richardson (1986; $R^2 = 0.81$ & 0.70 for unmasked and masked, respectively). The covariates of the model are shown in Fig. 16(a-d) and all covariates (*OHCr*, *IHCr*, *ENV*, & *AP*) were significant across all noise conditions ($p < 0.001$), meaning all of the simulated sources were important to appropriately model the CR at 45 Hz, 80 dB SPL. Figure 17 shows the real CR waveform and the model waveform for the unmasked condition (top) and the 7th masking condition (bottom). Morphologically, and especially in the unmasked condition, the fitted function did an excellent job of modeling the real data, including the distortions. In the masked

condition the model fit the data well morphologically, with the exception of one peak amplitude that was not quite fulfilled by the model (Fig. 17 bottom). The model fit and residuals are plotted in Fig. 18 for the unmasked (top) and 3700 Hz masking condition (bottom). The model fit shows a good correlation, and residuals are centered around zero indicative of a constant variance.

In order to examine the effect of all masking conditions on the model covariates, as opposed to just the unmasked and 3700 Hz masking condition as plotted in the previous figures, the estimated coefficients for each model covariate are plotted as a function of masking condition in Fig. 19. While the actual values of each covariate's coefficients may not be compared to one another, given that the initial amplitudes of the covariates going in to the model were somewhat arbitrary, it is interesting to compare the shape of the growth functions of the coefficients to each other, as well as the sign of the value and changes in the sign, as this may an indicator of phase and phase changes within each covariate. Figure 19(a) shows the growth function of the coefficients of the *OHCr* covariate. The values are negative, meaning a 180° phase shift to the *OHCr* was used to make the model. Additionally, the function grows (although negatively) in a fairly linear manner, which coincidentally, was what we predicted the response should do to a low-frequency stimulus such as 45 Hz if the response was mainly comprised of OHC (Chertoff et al., 2016): that the contribution of *OHCr* increases linearly as more of the cochlea becomes available to respond.

Figure 19(b & c) show the growth functions of the *IHCr* and *ENV* covariates, respectively. The growth function of the coefficients for *IHCr* has a sigmoidal growth function, similar to that of the CAF, meaning that the contribution of this covariate to the model changes nonlinearly as more of the cochlear partition becomes available to respond. Furthermore, unlike the *OHCr* function, the *IHCr* function changes sign twice, which, if coefficient sign can indeed be interpreted as phase, would mean there are two phase reversals in the *IHCr* as the cutoff frequency location of the masker moves from apex to base. The evidence for this in the literature and its implications in our data will be discussed in the next section. The *ENV* coefficient growth function remains positive in value and its shape can be described in two parts: a steep linear upward slope peaking at the 8th masking condition (6000 Hz cutoff), followed by a

shallower fairly-linear downward slope. The relation between this covariate and the data can be seen by comparing this growth function to how much this positive bias can be seen in the CR waveforms at each masking condition shown in Fig. 3. It is also important to note that the peak amplitude of *ENV* occurs at the same masking condition as the *IHCr* coefficient's maximum negative amplitude.

The coefficients of the phase-locked *AP* pulse train covariate (Fig. 19[d]) hover around zero for this stimulus condition (45 Hz, 80 dB SPL). This may be due to two possibilities: either the apical AN is not activated at such a low frequency and sound level or the covariate *AP* does not accurately simulate these action potentials.

The CAF of the low-frequency CR is sigmoidal, as shown by Kamerer et al. (2016) and schematized by the solid line in Fig. 1(a). The 45 Hz CAF is the amplitude of 45 Hz (F0) as a function of distance from the apex, which has been converted from the masker cutoff frequency. The CR model created in this experiment was used to create a model CAF. Only the *IHCr* and *OHCr* covariates were used because they are the only variables that contain F0. The 45 Hz component of each covariate was considered the complex number at 45 Hz peak in the FFT. The complex number containing both magnitude and phase was multiplied by the covariates' corresponding regression coefficient at each masker cutoff frequency. The magnitude and phase of *IHCr* and *OHCr* at each masking condition was linearly added and the magnitude of the sum was plotted as a function of distance from the apex corresponding to each masking condition. The resulting model CAF is shown with the real 45 Hz CR data in Fig. 20.

Experiment 4: Discussion

The purpose of this experiment was to model the low-frequency CR with the intent of identifying the most probable cellular sources to target in future physiologic experiments. Two statistical approaches worked from different directions to converge on similar conclusions regarding the sources that best model the 45 Hz, 80 dB SPL cochlear response. Independent component analysis worked backward trying to decompose the whole CR into its independent cellular sources, while the linear regression model worked forward by fitting simulated cellular sources together into a whole model CR.

The results of ICA were mixed. The FastICA algorithm worked near-perfectly on the simulated data. Though the mixture created by the simplified hair cell sources did not create a realistic model of the 45 Hz CR, there were similar properties –such as the same number of data points and a comparable spectrum –that would allow it to test ICA's ability to perform with mixtures like the CR. For this purpose, ICA was able to recover the sources that made up the two simulated mixtures. When the same algorithm was performed on the real 45 Hz cochlear response, ICA was able to recover two independent sources. One of the sources morphologically looked similar to the CR waveform itself, while the other looked like a noisy higher-frequency sinusoid. Temporally, these waveforms did not match either of the simulated hair cell responses. Spectrally, however, one source contained the fundamental frequency and the odd harmonics, much like the simulated OHC; the other source contained a large Env component and primarily even harmonics, similar to the simulated IHC response. While the second source morphologically did not seem to match that of an IHC, it did look similar to a response known to be generated by the AN –the auditory nerve overlapping waveform (ANOW).

At frequencies below 1 kHz, the CM approximates the stimulus and is sinusoidal. The apical auditory nerve response at these frequencies is phase-locked, or prefers to fire at a specific phase of the stimulus, usually to either the positive or negative phase of the CM. If the response is recorded again to a reversed-polarity stimulus, the neural response will remain locked to the same phase of the CM. If this stimulus polarity reversal is repeated many times and averaged, the CM will be canceled as it is 180° out-of-phase between polarities, and the neural response will remain as a sinusoidal waveform at twice the stimulus frequency (Henry, 1995; Lichtenhan et al., 2013, 2014). We did not expect the ANOW in our data, because we did not average across stimulus polarities. The ANOW is a fairly new electrophysiologic measure and currently has only been recorded at >300 Hz. Therefore, we cannot rule out the possibility that different populations of the AN may prefer to lock on to different phases of the CM response, or the possibility that because the stimulus frequency is low enough, half of a 45 Hz cycle is longer than the refractory period of an action potential (1-2 ms), so the apical AN may fire twice per cycle. This would generate energy at F2 even when recording with one stimulus polarity.

Thus, we were left with two possible explanations for the results of ICA on the 45 Hz CR. The first was that it worked well and truly recovered the independent cellular sources that make up the low-frequency CR. We speculate that if this is the case, these sources are likely OHC, and either IHC or apical auditory nerve. The second explanation is that ICA did not work well, likely because we violated one or more assumptions necessary for ICA to perform. One possible assumption violation is that the two (or more) sources of the CR are not statistically independent, i.e. the response from one source is correlated with another. This may be true, considering an OHC, IHC, and even an auditory nerve response is just that: a *response* to the same stimulus. Though an argument against this is that since the drives of excitation for these cells are different, then they aren't necessarily responding to the exact same stimulus. Evidence for this is the fact that ICA did work on the simulated hair cells, where the OHC response was derived from the stimulus and the IHC response was derived from the differentiated stimulus, because the IHC responds to the velocity of the BM. The second possible reason for an ICA failure was our attempt to manipulate the assumption that the number of sources must equal the number of electrodes. The CR was recorded using only one RW electrode. In order to recover two sources, we tried four methods of circumventing this assumption. First, we tried to use the CR in two different masking conditions as separate mixtures. FastICA would not work with these data. This is likely due to the addition of another variable, noise, to one of the recording situations. Second, we tried to use two different trials –the CR recorded in the same animal approximately 5 minutes apart –as two different mixtures. This also did not work in the FastICA algorithm, likely because the two trial mixtures were almost identical and therefore did not contain enough differences to be de-correlated in preprocessing and useful to ICA. Third, we attempted to partially uphold the assumption and recorded the CR consecutively from two locations on the cochlea: the RW niche and the bone surrounding the apex. In order to avoid inducing endolymphatic hydrops in the animal cochlea, we rested the electrode on the bone outside of the apical portion of cochlea instead of opening a hole in the bone where a better recording could be obtained. Unfortunately, this resulted in a CR amplitude that was too small and noisy relative to the RW recording for ICA to work. Our last attempt, reported in this study, was to use CRs recorded to two different stimulus polarities as

two mixtures. Using this method, FastICA did recover the sources reported in the results. Even though ICA was able to recover sources, it is hard to know whether this truly upheld the assumption. Thus, we were not able to conclude the validity of ICA on the 45 Hz CR.

To complement and bring clarity to the results of ICA, a more in-depth model was created and regressed to the real CR data to provide evidence for identifying the cellular origins of the low-frequency CR. The model simulated the most probable sources of the low-frequency CR as suggested by the results of Experiments 2 & 3 and the results of ICA. This included more realistic responses from hair cells, a stimulus envelope response (simulating the EFR, either from AN or hair cells), and simulated phase-locked apical AN responses. Linear regression analysis found all of these sources to be significant contributions to the model of the 45 Hz CR and these four sources were able to explain much of the real data. Additionally, the *IHCr* and *OHCr* covariates modeled the 45 Hz CAF fairly well. The apical region of the CAF was comparably flat, and there was a region of amplitude reduction in the model also seen in the real CAF. The modeled region of reduction did not quite align with that of the real data, suggesting the model is close but could be better. One way that the model could improve is to add several hair cell simulations to the model. Each hair cell covariate was simulated from real data recorded from hair cells at several different locations along the cochlear partition stimulated with different frequencies. It might behoove us to add more hair cell covariates to the model to simulated responses from different regions of the cochlea, e.g. basal, middle, and apical OHC & IHC covariates. This model could provide more location-specific estimates of hair cell responses and better model the CAF or the mid-cutoff-frequency masked CR.

Future physiologic studies are needed to confirm the identity of the suspected sources of the low-frequency cochlear response and to quantify each cell's contribution, but the regression model can be manipulated to estimate the contribution of the covariates to the CR waveform at each masking condition. The original regression model used covariates with somewhat arbitrary amplitudes. The hair cell covariates, *OHCr* and *IHCr*, had amplitudes determined by the transducer curves of real intracellular recordings, while the amplitudes of *ENV* and *AP* were nonrealistic values. To get a better idea of the

sources' relative contributions to the response, the covariates and the 45 Hz, 80 dB SPL CR were normalized to their maximum amplitudes (Fig. 21) so the max amplitude of each covariate and the CR was equal to one. The resulting estimated coefficients, then, could be compared across covariates (Fig. 22). Since the sign of the coefficient indicates phase, absolute values of the coefficients can be thought of as the contribution of each covariate relative to the max amplitude of the CR. The values of most interest are comparisons of the two simulated hair cell covariates. At the lowest masker cutoff frequency (146 Hz) the IHC response is twice as large as the OHC response. As masking cutoff frequency increases and more of the cochlea becomes available to respond, the OHC response starts to dominate the CR while the IHC response drops and reverses phase for several masking conditions. In the highest masker cutoff frequencies, the OHC and IHC responses are about equal. If this is true, then it contradicts the traditional belief that the OHC response is the dominating potential recorded in the cochlear microphonic. The pattern of *ENV* coefficients as a function of masking condition follow closely with the *IHCr* coefficients, suggesting that if the low-frequency EFR seen in the CR is indeed the SP, it is likely dominated by the inner hair cells.

DISCUSSION OF CELLULAR SOURCES OF THE LOW-FREQUENCY COCHLEAR RESPONSE

Outer Hair Cell

The plethora of evidence presented in the study for OHC as a primary source of the low-frequency CR comes as no surprise as it aligns with most of the literature showing that the response to a single-polarity tonal stimulus recorded at the RW is primarily CM –dominated by OHC responses. Experiment 1 showed that the distortion found in the low-frequency CR originated from nonlinearities in the cochlea. Outer hair cells driven at high stimulus levels to saturation will produce odd harmonics if their transducer curve is symmetric, and indeed there was significant amount of F3 and F5 in the low-frequency CR. In Experiment 2, the growth of the F3 harmonic was fairly small relative to the other two harmonic components and the shape of the curve relatively flat, meaning that F3 grew with a fairly similar pattern to that of F0. If the majority of F0 and F3 originated from the same physical process, i.e. current flow through OHC modified by the OHC transduction mechanism, then it makes sense that the ratio of the amplitude of these two components remains fairly constant as a function of masking condition. There was, however, a significant curvature to the F3/F0 growth curve, and one explanation for this is that the F0 component of the response originates not just from OHC but also from some other out-of-phase source which could reduce the peak amplitude of the F0 at certain masking conditions or spatial locations along the cochlear partition, increasing the F3/F0 ratio.

In Experiment 3, there was no significant effect of ouabain on the F3 component, though the F0 did show some changes dependent on treatment group and masking condition (or distance from the apex). Kamerer et al. (2016) did show some effect of ouabain on basal OHCs, manifested as ~5 dB reductions to 16 & 24 kHz DPOAEs. If the F0 was entirely comprised of the OHC response and OHCs were, in fact, affected by ouabain, then there should have also been changes to the F3 harmonic component. Indeed, there was a trend to the F3 growth curve that mimicked the F0, even if that change was not statistically significant. Another explanation is that the basal auditory nerve does contribute a small amount to the F0 but is out-of-phase and its removal resulted in the slight increase seen in the OHC response.

In Experiment 4, independent component analysis recovered two sources of the 45 Hz, 80 dB SPL CR. One of these sources contained odd harmonics (F3 & F5), suggesting the identity of this source as OHC. A simple simulation of an OHC based on the transducer function created from Russell, Cody, and Richardson (1986) recordings of OHC receptor potentials that also produced odd harmonics was used as a known source for a successful test of ICA. The analysis was able to pull this source back out of two mixtures containing the simulated OHC. Spectrally, the simulated OHC matched the recovered source, leading us to conclude that OHCs producing odd order harmonics were a significant source of the 45 Hz CR. This was confirmed by the regression analysis, which found that a (more realistic) simulated OHC was a significant covariate in modeling the CR. In fact, the model fit was dependent on the real data from whence it as derived. Three possible OHC models were tried in the model from data published by three groups who reported transducer functions for OHCs using different intracellular recording methods, and maybe more importantly, different characteristic-frequency cells and stimulus frequencies. A result of these recording differences led to three transducer curves with different features such as symmetry, steepness, and sharpness of saturation. The simulated OHC that fit the 45 Hz CR model best was Dallos et al. (1986) which, of the three, had the steepest linear phase and sharpest compression phase; and was asymmetric, with a depolarizing phase larger than the hyperpolarizing phase. This transducer curve was recorded from a third turn OHC stimulated with an 800 Hz tone, so the characteristic frequency of the cell was close to the stimulus frequency. This may be evidence of the spatial location of the OHCs responding to the low-frequency CR –in that these are apical cells with a characteristic frequency around that of the stimulus. This may be taken as confirmation that the masking noise is suppressing basal OHC responses, however, the Russell, Cody, & Richardson (1986) transducer function which was recorded from a basal OHC stimulated with a low-frequency tone also fit the data well, and the apical OHC transducer function fit the data even in the unmasked condition where we would expect basal contributions. Additionally, the estimated coefficients for the *OHCr* were negative, meaning that the simulated OHC was best fit to the model when the waveform was flipped 180°. This would also reverse the DC response from this cell from positive to negative, such that the hyperpolarization was larger than the depolarization. Indeed, if one

looks at the unmasked low-frequency CR, especially seen in the 85 Hz, 90 dB SPL condition (Fig. 6, boxed panel), one would notice a DC shift in the negative direction. The outer hair cell DC response to frequencies below the characteristic frequency of the cell at moderate sound levels is dominated by a hyperpolarizing phase (or negative DC) which could explain the negative coefficients, however, at high stimulus levels the OHC is shown to have a positive DC. Due to all these possibilities, it may be overzealous to draw conclusions regarding the location of the OHCs responding to the 45 Hz tone.

One hypothesis of Kamerer et al. (2016) was that low-frequency tones would elicit a CAF that grew more linearly than that of higher-frequency tones, on the basis that for every millimeter of the cochlear partition that became unmasked, a purely OHC response should grow a fairly equal amount. On the contrary, they found that the 45 Hz CAF was similar to the 762 Hz CAF in that it grew in a sigmoidal shape, with shallow growth in the apex and even reduction in amplitude at some masking conditions. The linear regression analysis in this study revealed that the *OHCr* covariate coefficients grew (in absolute value) fairly linearly indeed as a function of masking condition. If the F0 of the low-frequency CR was comprised of entirely OHC response then the CAF may, in fact, have grown in a more linear fashion. Our hypothesis was that a(n) additional cellular source(s) was contributing to the F0 but out-of-phase with that of the OHC response, and that this source may grow nonlinearly such that it comes close to equaling the OHC contribution where the CAF reduces in amplitude.

Inner Hair Cell

The results of Experiments 1 & 2 showed a significant even harmonic contribution to the low-frequency CR, not seen at higher frequencies. At higher frequencies, the CR is dominated by OHC responses (CM) and basal AN. There were several possible sources of the even-order harmonics in the CR: OHCs, IHCs, and the apical AN. The linear regression model provided the strongest evidence for this source to be IHC. The potential for this source to be entirely made of OHC or apical AN can be ruled out for several reasons:

Outer hair cells can produce even-order harmonic distortion when they have asymmetric transducer curves. This can occur as a shift in the operating point (or the output of the cell at rest, i.e.

when there is 0 Pa input) of the curve, which is known to happen when the endocochlear potential is altered, for example, in the case of endolymphatic hydrops (Salt et al., 2009; Brown et al., 2009). Endolymphatic hydrops can be induced by invasive surgeries involving opening the bony portion of the cochlea and even by low-frequency tone exposure (Kirk & Patuzzi, 1997; Salt, 2004). The present study recorded the CR from only the RW niche, so we do not expect to have surgically caused endolymphatic hydrops, however, the use of low-frequency stimuli may have produced a shift in the OHC operating point. So, which hair cell produced the even-order harmonics? The linear regression model which used an IHC response as the source of the even harmonic distortion added significantly to a model which explained a high proportion of the 45 Hz CR waveform in a way that even-order harmonics produced by an OHC could not. Even if both the OHC and IHC transducer functions were asymmetric and produced even harmonics –which may be true given that an asymmetric OHC function fit the 45 Hz CR the best of the three OHC functions –there is a phase difference between the two hair cells due to their mechanisms of excitation that is important to the model. Inner hair cells are driven at low frequencies primarily by drag of the viscous fluid in the reticular lamina (RL) –tectorial membrane (TM) gap (see *Appendix A: Hair Cell Transduction*) that means that they respond to stimulus acceleration (or BM velocity) as opposed to OHCs which are mechanically coupled to the TM and therefore respond to stimulus velocity (or BM displacement). This results in a 90° phase difference between the two hair cell responses that precedes the generation of the receptor potential modulated by each cell's transducer curve, and this phase difference is important to the model. There is no group delay to the CM, meaning that we should not expect there to be a time delay in our CR recordings from two populations of hair cells residing at different locations along the cochlear partition, therefore, we should not expect a phase difference from any two populations of OHCs at different locations (He et al., 2012).

The apical AN did contribute to the low-frequency CR, and likely did contribute to both the even and odd-order harmonics. The contribution of the apical AN to the low-frequency CR will be discussed in the next section, however, the *AP* covariate of the linear regression model could not explain all of the F2 component. A source covariate that could explain most of the even harmonics was an IHC source. The

linear regression model used a simulated IHC response using transducer curves from intracellular recordings solved using the second derivative of the stimulus to imitate the IHC response to stimulus acceleration. The resulting simulated response waveform from an IHC shows a large amount of distortion in addition to a 90° phase lead. Both the negative and positive peaks of the waveform are split. Guinan (2012) modeled how a non-perfect sine wave stimulus and the combination excitatory/inhibitory mechanisms of the IHC (see *Appendix A: Hair Cell Transduction*) could result in peak-splitting of the response at low-frequencies, which may be reflected in the AN response like those recorded by (Nelson Yuan-sheng Kiang, 1990). We have shown in Experiment 1 that our toneburst stimuli are not perfect, and a response to acceleration would exaggerate even small distortions, resulting in the split-peak response waveform of the *IHCr* covariate. As it turns out, these extra deflections are important to modeling all the deflections in the 45 Hz CR waveform, providing evidence for a significant source to be IHC.

An IHC response would also contribute to the F0, or the CAF, only 90° out-of-phase with the OHC. The linear regression model shows that the *IHCr* covariate grows nonlinearly as a function of masking condition, and its peak negative amplitude aligns fairly well with where the CAF reduces in amplitude. In fact, summing the *OHCr* and *IHCr* F0 components results in a sigmoidal CAF similar to that of the real CAF data, with a reduction in amplitude at certain masking conditions. The nonlinear growth of the *IHCr* coefficients including sign-changes, which could be considered phase-reversals, have been shown in chinchilla AN responses to low-frequencies, which Guinan argues is a reflection of phase-reversals in the IHC receptor potential. Guinan presents a model in which the dominating excitatory mechanism of apical IHCs to low-frequency tones is the typical shearing motion of the RL-TM, while basal IHC responses to low-frequencies at high levels (such as those used in this study) may be dominated by the cilia-slant mechanism (see *Appendix A*) resulting in phase variances between populations of IHCs. So, while apical IHCs may be out-of-phase with OHCs and reduce or halt growth of the CAF, basal IHCs respond in-phase with OHCs, contributing to the growth of the CAF curve at higher masking cutoff frequencies. Additionally, most simulations of the IHC responses, i.e. those used in this study, would show contribution to the F3, as an asymmetric transducer function driven to its compressive

phase will produce all higher-order harmonics. The IHC used in the regression analysis for Experiment 4 produced an F3 component (Fig. 12, FFT) though not as large as the F3 produced by an OHC.

While the AC response from IHCs has not been well-recorded in extracellular potentials like the CR, the DC response of IHCs is largely attributed to the SP recorded in ECochG (Zheng et al., 1997). At low stimulus levels and frequencies, the IHC produces a large depolarizing bias relative to hyperpolarization, resulting in a positive DC component, while OHCs produce a larger hyperpolarizing phase resulting in a negative DC component (see *Appendix A: Hair Cell Transduction*). The *ENV* covariate which simulates a large positive shift in the CR data at mid-masker-cutoff frequencies, would align with an IHC source. However, there is some evidence that this Env component may be generated by the auditory nerve as well as both inner and outert hair cell. First the Env component was slightly affected in the mid-basal regions of the cochlea where ouabain resulted in the largest loss of the auditory nerve. Additionally, the growth of the Env/F0 was significantly different from both the higher order harmonics (F2 & F3) at all stimulus conditions, suggesting that the source(s) which produces the Env cannot be identified as either source that produces the F2 or F3. Additionally, the OHC can contribute a significant amount to the SP (Durrant et al., 1998). Therefore, the Env distortion-product found in the low-frequency CR may be a combination of auditory nerve and both inner and outer hair cell contributions.

Apical Auditory Nerve as a Source of the Low-Frequency Cochlear Response

The low-frequency CR at both 45 & 85 Hz contained phase-locked action potentials on both positive and negative slopes of the waveform. High-pass filtering of the waveform revealed that there may be even more action potentials at other phases of the waveform as well, though much smaller and wider, indicative of summed responses from smaller neural populations and/or are less phase-locked. The sinusoidal shape of the ANOW response to higher frequency tones (~300 Hz) is effectively a peristimulus time histogram of the neural response to a sinusoidal tone. At lower frequencies, like 45 Hz, the AN may be able to phase-lock more accurately because the refractory period of an AN fiber is much smaller than the period of the stimulus. This could result in the pulse-train seen in the high-pass filtered waveform. A pulse-train will generate spectral energy at 1/T, or the period between pulses, which is the same as either

F0 following one phase of the waveform (such as AP+) or 2F0 following two phases of the waveform (AP+ and AP-); and equal energy at the intervals n/T, or all harmonics. In the case of 45 Hz, a pulse-train locked to one phase of each cycle results in equal energy at 45, 90, 135 Hz, and all higher order harmonics. In the model of a pulse-train with APs on both the positive and negative slopes of the CR waveform, an FFT showed energy at 90, 180, 270 Hz, and higher order harmonics, however, not near the amount of F2 seen in the CR waveform. In fact, even when the simulated AP pulse-train was made equal in amplitude to the CR (which is not ever seen in the real data) the amount of harmonic energy produced by the pulse-train could not equal that seen in the real CR data.

In addition to the phase-locked response, the auditory nerve may be contributing an EFR that comprises part of the Env distortion-component found in the low-frequency CR. It is difficult to separate any apical auditory nerve contribution from the hair cell contributions at this point. We know that ouabain damaged the AN such that there were no CAP responses above 8 kHz, and there was significant reduction in amplitude to the CAP response down to 1 kHz (Kamerer et al., 2016). However, we do see phase-locked action potentials in the waveform and therefore know that the entire AN was not obliterated by ouabain. We can conclude that at least the apical AN is indeed contributing to the low-frequency CR, but is not the sole source of the higher-order harmonics.

Future Directions

This study presents evidence for several cellular sources of the low-frequency cochlear response and evidence that these contributors are likely outer and inner hair cells and the apical auditory nerve. Additionally, the statistical analyses in Experiment 4 may have separated the individual contributions to the waveform which could be useful in eventually transforming the low-frequency CR into a diagnostic tool. In order to definitively prove each cell's contribution, additional physiologic experiments like that of Experiment 3 must be performed. The current study will be followed by a chronic ouabain experiment, much like that of Kamerer et al. (2016) and Experiment 3, except with a longer application of ouabain such that the entire AN is damaged. This chronic application of ouabain will be followed by recordings of the low-frequency CR and observation of effects on the harmonics and the morphology of the waveform.

We would expect a loss of all phase-locked APs and some effect on the higher-order harmonics. Once the contribution of the AN is quantified and successfully modeled, the next experiment would be to damage inner hair cells. The systematic and specific damage to IHCs without affecting OHCs is more difficult, as much of their anatomy is shared, however, there have been several successful methods of eliminating the IHC response without affecting OHCs. Carboplatin-treated chinchillas have been used to damage IHCs and study auditory neuropathy (Cowper-Smith et al., 2010). A mouse model has also been used to knock-out IHCs; slc19a2-null mice lose inner hair cells when feed a thiamine-poor diet. Fleming et al. (2001; 2003; 1999) created a genetic strain of mice that lacked the SLC19A2 gene. This gene codes for a thiamine receptor found throughout the body, but within the cochlea, specific to inner hair cells. In 2006, Liberman and colleagues were the first hearing scientists to utilize these mice as a model for auditory neuropathy. They found that feeding these slc19a2-null mice a diet low in thiamine would systematically destroy IHCs along the length of the cochlear partition in a matter of weeks, with minimal damage to outer hair cells. In a future study comparison of the CR and the simulated source components of the mice with and without IHC damage could prove that 1) IHCs are a significant source of the low-frequency CR and 2) the potential of the low-frequency CR to assess the health of IHCs.

CONCLUSIONS

1. The low-frequency cochlear response contains harmonic distortions that are a product of nonlinear mechanisms in the cochlea.
2. Independent component analysis recovered two sources of the low-frequency cochlear response and therefore shows promise as a method of analyzing the components of the cochlear response recorded from the round-window, once the approach is refined and assumptions met.
3. Linear regression analysis was able to model the low-frequency cochlear response using a simulated outer hair cell, inner hair cell, stimulus envelope response, and auditory nerve action potentials.
4. The basal auditory nerve is likely not a significant cellular contributor to the low-frequency cochlear response.
5. The apical auditory nerve is one cellular source of the low-frequency cochlear response.
6. Outer hair cells are a significant cellular source of the low-frequency cochlear response.
7. Inner hair cells are likely a significant cellular source of the low-frequency cochlear response.

REFERENCES

Adrian, E. D., Bronk, D. W., & Phillips, G. (1931). The nervous origin of the Wever and Bray effect. *The Journal of Physiology*, *73*, 2P–3.

Allen, J. B. (1980). Cochlear micromechanics—A physical model of transduction. *The Journal of the Acoustical Society of America*, *68*(6), 1660–1670. http://doi.org/10.1121/1.385198

Boettcher, F. A., Mills, J. H., Dubno, J. R., & Schmiedt, R. A. (1995). Masking of auditory brainstem responses in young and aged gerbils. *Hearing Research*, *89*(1-2), 1–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8600113

Bourien, J., Tang, Y., Batrel, C., Huet, a., Lenoir, M., Ladrech, S., … Wang, J. (2014). Contribution of auditory nerve fibers to compound action potential of the auditory nerve. *Journal of Neurophysiology*, *112*(5), 1025–1039. http://doi.org/10.1152/jn.00738.2013

Brown, D. J., Hartsock, J. J., Gill, R. M., Fitzgerald, H. E., & Salt, A. N. (2009). Estimating the operating point of the cochlear transducer using low-frequency biased distortion products. *The Journal of the Acoustical Society of America*, *125*(4), 2129–45. http://doi.org/10.1121/1.3083228

Brownell, W. E. (1984). Microscopic observation of cochlear hair cell motility. *Scanning Electron Microscopy*, (Pt 3), 1401–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6505622

Cardoso, J. F., & Souloumiac, A. (n.d.). An efficient technique for the blind separation of complex sources. In *[1993 Proceedings] IEEE Signal Processing Workshop on Higher-Order Statistics* (pp. 275–279). IEEE. http://doi.org/10.1109/HOST.1993.264552

Chan, D. K., & Hudspeth, A. J. (2005). Mechanical Responses of the Organ of Corti to Acoustic and Electrical Stimulation In Vitro. *Biophysical Journal*, *89*(6), 4382–4395. http://doi.org/10.1529/biophysj.105.070474

Charaziak, K. K., Shera, C. A., & Siegel, J. H. (2017). Using cochlear microphonic potentials to localize peripheral hearing loss. *Frontiers in Neuroscience*, *11*(APR). http://doi.org/10.3389/fnins.2017.00169

Cheatham, M. a, & Dallos, P. (1999). Response phase: a view from the inner hair cell. *The Journal of the Acoustical Society of America*, *105*(February 1999), 799–810. http://doi.org/10.1121/1.426269

Cheatham, M. A., & Dallos, P. (1994). Stimulus biasing: A comparison between cochlear hair cell and organ of corti response patterns. *Hearing Research*, *75*(1-2), 103–113. http://doi.org/10.1016/0378-5955(94)90061-2

Cheatham, M. A., Naik, K., & Dallos, P. (2011). Using the cochlear microphonic as a tool to evaluate cochlear function in mouse models of hearing. *Journal of the Association for Research in Otolaryngology : JARO*, *12*(1), 113–25. http://doi.org/10.1007/s10162-010-0240-5

Chertoff, M. E., Earl, B. R., Diaz, F. J., & Sorensen, J. L. (2012). Analysis of the cochlear microphonic to a low-frequency tone embedded in filtered noise. *The Journal of the Acoustical Society of America*, *132*(5), 3351–62. http://doi.org/10.1121/1.4757746

Chertoff, M. E., Earl, B. R., Diaz, F. J., Sorensen, J. L., Thomas, M. L. a, Kamerer, A. M., & Peppi, M. (2014). Predicting the location of missing outer hair cells using the electrical signal recorded at the round window. *The Journal of the Acoustical Society of America*, *136*(3), 1212. http://doi.org/10.1121/1.4890641

Chertoff, M. E., Kamerer, A. M., Peppi, M., & Lichtenhan, J. T. (2015). An analysis of cochlear response harmonics: Contribution of neural excitation. *The Journal of the Acoustical Society of America*, *138*(5), 2957–2963. http://doi.org/10.1121/1.4934556

Cody, A. R., & Russell, I. J. (1987). The response of hair cells in the basal turn of the guinea-pig cochlea to tones. *The Journal of Physiology*, *383*, 551–69. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1183089&tool=pmcentrez&rendertype=abstract\nhttps://www.dropbox.com/s/tq4fdaed0wt33zd/1987 - Cody, Russell - The response of hair cells in the basal turn of the guinea-pig cochlea to tones.pdf?dl

Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing*, *36*(3), 287–314. http://doi.org/10.1016/0165-1684(94)90029-9

Conijn, E. A., Brocaar, M. P., van Zanten, G. A., & van der Drift, J. F. (1992). Comparison between the

frequency specificities of auditory brainstem response thresholds to clicks with and without high-pass masking noise. *Audiology : Official Organ of the International Society of Audiology*, *31*(5), 284–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1449431

Cowper-Smith, C. D., Dingle, R. N., Guo, Y., Burkard, R., & Phillips, D. P. (2010). Synchronous auditory nerve activity in the carboplatin-chinchilla model of auditory neuropathy. *The Journal of the Acoustical Society of America*, *128*(1), EL56–62. http://doi.org/10.1121/1.3453764

Dallos, P. (1969). Comments on the Differential-Electrode HE content of this paper is probably well known. *J. Acoust. Soc. Am*, *45*(June), 999–1007.

Dallos, P. (1973). Basic Mechanisms in Hearing. In A. Møller (Ed.), *Basic Mechanisms in Hearing* (pp. 335–372). New York: Academic Press.

Dallos, P. (1985). Response characteristics of mammalian cochlear hair cells. *Journal of Neuroscience*, *5*(6), 1591–1608.

Dallos, P. (1986). Neurobiology of cochlear inner and outer hair cells: intracellular recordings. *Hearing Research*, *22*, 185–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3733539

Dallos, P., & Cheatham, M. A. (1976). Production of cochlear potentials by inner and outer hair cells. *The Journal of the Acoustical Society of America*, *60*(2), 510–2. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/993471

Dallos, P., & Cheatham, M. A. (1989). Nonlinearities in cochlear receptor potentials and their origins. *The Journal of the Acoustical Society of America*, *86*(5), 1790–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2808928

Dallos, P., & Durrant, J. D. (1972). On the derivative relationship between stapes movement and cochlear microphonic. *The Journal of the Acoustical Society of America*, *52*(4), 1263–1265. http://doi.org/10.1121/1.1913241

Dallos, P., & Sweetman, R. H. (1969). Distribution pattern of cochlear harmonics. *The Journal of the Acoustical Society of America*, *45*(1), 37–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5797145

Davis, H. (1968). Mechanisms of the Inner Ear. *Annals of Otology, Rhinology & Laryngology*, *77*(4), 644–655. http://doi.org/10.1177/000348946807700405

Davis, H., Deathrage, B. H., Eldredge, D. H., & Smith, C. A. (1958). Summating potentials of the cochlea. *The American Journal of Physiology*, *195*(2), 251–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13583157

Davis, H., Derbyshire, A. J., Lurie, M. H., & Saul, L. J. (1934). The electrical response of the cochlea. *American Journal of Physiology*, *107*, 311–332.

de Araujo, D. B., Barros, A. K., Estombelo-Montesco, C., Zhao, H., da Silva Filho, A. C. R., Baffa, O., … Ohnishi, N. (2005). Fetal source extraction from magnetocardiographic recordings by dependent component analysis. *Physics in Medicine and Biology*, *50*(19), 4457–4464. http://doi.org/10.1088/0031-9155/50/19/002

Dolphin, W. F., & Mountain, D. C. (1991). Scalp potentials follow the low frequency envelope of complex acoustic stimuli. In *Proceedings of the 1991 IEEE Seventeenth Annual Northeast Bioengineering Conference* (pp. 215–216). IEEE. http://doi.org/10.1109/NEBC.1991.154651

Dolphin, W. F., & Mountain, D. C. (1992). The envelope following response: Scalp potentials elicited in the mongolian gerbil using sinusoidally AM acoustic signals. *Hearing Research*, *58*(1), 70–78. http://doi.org/10.1016/0378-5955(92)90010-K

Domenica Karavitaki, K., & Mountain, D. C. (2007). Imaging Electrically Evoked Micromechanical Motion within the Organ of Corti of the Excised Gerbil Cochlea. *Biophysical Journal*, *92*(9), 3294–3316. http://doi.org/10.1529/biophysj.106.083634

Durrant, J. D., Wang, J., Ding, D. L., & Salvi, R. J. (1998). Are inner or outer hair cells the source of summating potentials recorded from the round window? *The Journal of the Acoustical Society of America*, *104*(1), 370–377. http://doi.org/10.1121/1.423293

Earl, B. R., & Chertoff, M. E. (2012). Mapping auditory nerve firing density using high-level compound action potentials and high-pass noise masking. *The Journal of the Acoustical Society of America*, *131*(1), 337–52. http://doi.org/10.1121/1.3664052

Eggermont, J. J. (1976). Analysis of compound action potential responses to tone bursts in the human and guinea pig cochlea. *The Journal of the Acoustical Society of America*, *60*(5), 1132–1139. http://doi.org/10.1121/1.381214

Elberling, C. (1974). Action Potentirals Along the Cochlear Partition Recorded from the ear Canal in Man. *Scandinavian Audiology*, *3*(1), 13–19. http://doi.org/10.3109/01050397409044959

Engebretson, A. M., & Eldredge, D. H. (1968). Model for the Nonlinear Characteristics of Cochlear Potentials. *The Journal of the Acoustical Society of America*, *44*(2), 548–554. http://doi.org/10.1121/1.1911119

Evans, E. F., & Elberling, C. (1982). Location-specific components of the gross cochlear action potential: an assessment of the validity of the high-pass masking technique by cochlear nerve fibre recording in the cat. *Audiology : Official Organ of the International Society of Audiology*, *21*(3), 204–227. http://doi.org/10.3109/00206098209072740

Fleming, J. C., Steinkamp, M. P., Kawatsuji, R., Tartaglini, E., Pinkus, J. L., Pinkus, G. S., … Neufeld, E. J. (2001). Characterization of a murine high-affinity thiamine transporter, Slc19a2. *Molecular Genetics and Metabolism*, *74*(1-2), 273–280. http://doi.org/10.1006/mgme.2001.3241

Fleming, J. C., Tartaglini, E., Kawatsuji, R., Yao, D., Fujiwara, Y., Bednarski, J. J., … Neufeld, E. J. (2003). Male infertility and thiamine-dependent erythroid hypoplasia in mice lacking thiamine transporter Slc19a2. *Molecular Genetics and Metabolism*, *80*(1-2), 234–241. http://doi.org/10.1016/S1096-7192(03)00141-0

Fleming, J. C., Tartaglini, E., Steinkamp, M. P., Schorderet, D. F., Cohen, N., & Neufeld, E. J. (1999). The gene mutated in thiamine-responsive anaemia with diabetes and deafness (TRMA) encodes a functional thiamine transporter, *22*(july), 305–308.

Freeman, D. M., & Weiss, T. F. (1990a). Hydrodynamic forces on hair bundles at high frequencies. *Hearing Research*, *48*(1-2), 31–36. http://doi.org/10.1016/0378-5955(90)90197-W

Freeman, D. M., & Weiss, T. F. (1990b). Superposition of hydrodynamic forces on a hair bundle. *Hearing Research*, *48*(1-2), 1–15. http://doi.org/10.1016/0378-5955(90)90195-U

Gold, T. (1989). Historical Background to the Proposal, 40 Years Ago, of an Active Model for Cochlear Frequency Analysis. In J. P. Wilson & D. T. Kemp (Eds.), *Cochlear Mechanisms: Structure, Function, and Models* (pp. 299–305). Boston, MA: Springer US. http://doi.org/10.1007/978-1-4684-5640-0_34

Guinan, J. J. (2012). How are inner hair cells stimulated ? Evidence for multiple mechanical drives. *Hearing Research*, *292*(1-2), 35–50. http://doi.org/10.1016/j.heares.2012.08.005

Hallpike, C. S., & Rawdon-Smith, A. F. (1934). The "Wever and Bray phenomenon." A study of the electrical response in the cochlea with especial reference to its origin. *The Journal of Physiology*, *81*(3), 395–408. http://doi.org/10.1113/jphysiol.1934.sp003143

He, W., Porsov, E., Kemp, D., Nuttall, A. L., & Ren, T. (2012). The group delay and suppression pattern of the cochlear microphonic potential recorded at the round window. *PloS One*, *7*(3), e34356. http://doi.org/10.1371/journal.pone.0034356

Henry, K. R. (1995). HSIRIrlG RES [ARCH Auditory nerve neurophonic recorded from the round window of the Mongolian gerbil, *90*, 176–184.

Honrubia, V., & Ward, P. (1970). Mechanism of production of cochlear microphonics. *The Journal of the Acoustical Society of …*, *47*(2). Retrieved from http://scitation.aip.org/content/asa/journal/jasa/47/2B/10.1121/1.1911920

Honrubia, V., & Ward, P. H. (1969). Dependence of the cochlear microphonics and the summating potential on the endocochlear potential. *The Journal of the Acoustical Society of America*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5804109

Hudspeth, a J., & Corey, D. P. (1977). Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(6), 2407–2411. http://doi.org/10.1073/pnas.74.6.2407

Hunyadi, B., Tousseyn, S., Mijović, B., Dupont, P., Van Huffel, S., Van Paesschen, W., & De Vos, M. (2013). ICA extracts epileptic sources from fMRI in EEG-negative patients: A retrospective validation study. *PLoS ONE*, *8*(11). http://doi.org/10.1371/journal.pone.0078796

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, *10*(3), 626–34. http://doi.org/10.1109/72.761722

Hyvärinen, A., & Oja, E. (1997). A Fast Fixed-Point Algorithm for Independent Componente Analysis.

James, C. J., & Lowe, D. (2003). Extracting multisource brain activity from a single electromagnetic channel. *Artificial Intelligence in Medicine*, *28*(1), 89–104. http://doi.org/10.1016/S0933-3657(03)00037-X

Jia, S., Dallos, P., & He, D. Z. Z. (2007). Mechanoelectric transduction of adult inner hair cells. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(5), 1006–1014. http://doi.org/10.1523/JNEUROSCI.5452-06.2007

Johnstone, B. M., & Boyle, A. J. (1967). Basilar membrane vibration examined with the Mössbauer technique. *Science (New York, N.Y.)*, *158*(3799), 389–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6061893

Johnstone, B. M., Johnstone, J. R., & Pugsley, I. D. (1966). Membrane Resistance in Endolymphatic Walls of the First Turn of the Guinea-Pig Cochlea. *The Journal of the Acoustical Society of America*, *40*(6), 1398–1404. http://doi.org/10.1121/1.1910239

Johnstone, B. M., Taylor, K. J., & Boyle, A. J. (1970). Mechanics of the guinea pig colea. *The Journal of the Acoustical Society of America*, *47*(2), 504–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5439649

Johnstone, B. M., & Yates, G. K. (1974). Basilar membrane tuning curves in the guinea pig. *The Journal of the Acoustical Society of America*, *55*(3), 584–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4206680

Kamerer, A. M., Diaz, F. J., Peppi, M., & Chertoff, M. E. (2016). The potential use of low-frequency tones to locate regions of outer hair cell loss. *Hearing Research*, *342*, 39–47. http://doi.org/10.1016/j.heares.2016.09.006

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *The Journal*

of the Acoustical Society of America, *64*(5), 1386–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/744838

Kiang, N. Y. (1965). *Discharge patterns of single fibers in the cat's auditory nerve. M.I.T. Press.* Cambridge, Mass.

Kirk, D., & Patuzzi, R. (1997). Transient changes in cochlear potentials and DPOAEs after low-frequency tones: the "two-minute bounce"revisited. *Hearing Research*, *112*, 49–68. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378595597001056

Klein, A. J. (1986). Masking effects on ABR waves I and V in infants and adults. *The Journal of the Acoustical Society of America*, *79*(3), 755–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3958315

Kokkinakis, K., & Loizou, P. C. (2008). Using blind source separation techniques to improve speech recognition in bilateral cochlear implant patients. *The Journal of the Acoustical Society of America*, *123*(4), 2379–2390. http://doi.org/10.1121/1.2839887

Kuwada, S., Andersont, J. S., Batra, R., Fitzpatrick, D. C., Teisser, N., & D'Angelo, W. R. (2002). Sources of the Scalp-Recorded Amplitude- Modulation Following Response. *Journal of American Academy of Audiology*, *13*, 188–204.

Landegger, L. D., Psaltis, D., & Stankovic, K. M. (2016). Human audiometric thresholds do not predict specific cellular damage in the inner ear. *Hearing Research*, *335*, 83–93. http://doi.org/10.1016/j.heares.2016.02.018

Lang, H., Schulte, B. a, & Schmiedt, R. a. (2005). Ouabain induces apoptotic cell death in type I spiral ganglion neurons, but not type II neurons. *Journal of the Association for Research in Otolaryngology : JARO*, *6*(1), 63–74. http://doi.org/10.1007/s10162-004-5021-6

Li, G., & Zhang, J. (1998). Sphering and Its Properties. *Sankhyā: The Indian Journal of Statistics, Series A*, *60*(1), 119–133. Retrieved from http://www.jstor.org/stable/25051187\npapers://1aa359f0-a4fa-4e9a-9435-15fa577db1cc/Paper/p2873\nhttp://www.jstor.org/stable/pdfplus/25051187.pdf

Liberman, M. C., Tartaglini, E., Fleming, J. C., & Neufeld, E. J. (2006). Deletion of SLC19A2, the high

affinity thiamine transporter, causes selective inner hair cell loss and an auditory neuropathy phenotype. *JARO - Journal of the Association for Research in Otolaryngology*, *7*(3), 211–217. http://doi.org/10.1007/s10162-006-0035-x

Lichtenhan, J. T., Cooper, N. P., & Guinan, J. J. (2013). A new auditory threshold estimation technique for low frequencies: proof of concept. *Ear and Hearing*, *34*(1), 42–51. http://doi.org/10.1097/AUD.0b013e31825f9bd3

Lichtenhan, J. T., Hartsock, J. J., Gill, R. M., Guinan, J. J., & Salt, a N. (2014). The auditory nerve overlapped waveform (ANOW) originates in the cochlear apex. *Journal of the Association for Research in Otolaryngology : JARO*, *15*(3), 395–411. http://doi.org/10.1007/s10162-014-0447-y

Lindsen, J. P., & Bhattacharya, J. (2010). Correction of blink artifacts using independent component analysis and empirical mode decomposition. *Psychophysiology*, *47*(5), 955–960. http://doi.org/10.1111/j.1469-8986.2010.00995.x

McGuirt, J. P., & Schulte, B. a. (1994). Distribution of immunoreactive alpha- and beta-subunit isoforms of Na,K-ATPase in the gerbil inner ear. *The Journal of Histochemistry and Cytochemistry : Official Journal of the Histochemistry Society*, *42*(7), 843–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8014467

McLean, W. J., Smith, K. A., Glowatzki, E., & Pyott, S. J. (2009). Distribution of the Na,K-ATPase α Subunit in the Rat Spiral Ganglion and Organ of Corti. *Journal of the Association for Research in Otolaryngology*, *10*(1), 37–49. http://doi.org/10.1007/s10162-008-0152-9

Naunton, R. F., & Zerlin, S. (1976). Human whole-nerve responses to clicks of various frequency. *Audiology*, *15*, 1–9.

Nelson Yuan-sheng Kiang. (1990). Curious oddments of auditory-nerve studies. *Hearing Research*, *49*(1-3), 1–16. http://doi.org/10.1016/0378-5955(90)90091-3

Nowotny, M., & Gummer, A. W. (2006). Nanomechanics of the subtectorial space caused by electromechanics of cochlear outer hair cells. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2120–5. http://doi.org/10.1073/pnas.0511125103

Nowotny, M., & Gummer, A. W. (2011). Vibration responses of the organ of Corti and the tectorial membrane to electrical stimulation. *The Journal of the Acoustical Society of America*, *130*(6), 3852–72. http://doi.org/10.1121/1.3651822

Nuttall, A. L., Brown, M. C., Masta, R. I., & Lawrence, M. (1981). Inner hair cell responses to the velocity of basilar membrane motion in the guinea pig. *Brain Research*, *211*(1), 171–174. http://doi.org/10.1016/0378-5955(80)90080-5

O'Brien, W. J., Lingrel, J. B., & Wallick, E. T. (1994). Ouabain Binding Kinetics of the Rat Alpha Two and Alpha Three Isoforms of the Sodium-Potassium Adenosine Triphosphate. *Archives of Biochemistry and Biophysics*, *310*(1), 32–39. http://doi.org/10.1006/abbi.1994.1136

Oates, P., & Stapells, D. R. (1997). Frequency specificity of the human auditory brainstem and middle latency responses to brief tones. I. High-pass noise masking. *The Journal of the Acoustical Society of America*, *102*(6), 3597–3608. http://doi.org/10.1121/1.420148

Palmer, A. R., & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing Research*, *24*(1), 1–15. http://doi.org/10.1016/0378-5955(86)90002-X

Patuzzi, R. B., & Yates, G. K. (1987). The low-frequency response of inner hair cells in the guinea pig cochlea: implications for fluid coupling and resonance of the stereocilia. *Hearing Research*, *30*(1), 83–98. http://doi.org/10.1016/0378-5955(87)90186-9

Patuzzi, R. B., Yates, G. K., & Johnstone, B. M. (1989). The origin of the low-frequency microphonic in the first cochlear turn of guinea-pig. *Hearing Research*, *39*(1-2), 177–188. http://doi.org/10.1016/0378-5955(89)90089-0

Patuzzi, R., & Sellick, P. M. (1983). A comparison between basilar membrane and inner hair cell receptor potential input-output functions in the guinea pig cochlea. *The Journal of the Acoustical Society of America*, *74*(6), 1734–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6655131

Pierre, S. V, Sottejeau, Y., Gourbeau, J.-M., Sánchez, G., Shidyak, A., & Blanco, G. (2008). Isoform specificity of Na-K-ATPase-mediated ouabain signaling. *American Journal of Physiology. Renal*

Physiology, *294*(4), F859–66. http://doi.org/10.1152/ajprenal.00089.2007

Ponton, C. W., Don, M., & Eggermont, J. J. (1992). Place-specific derived cochlear microphonics from human ears. *Scandinavian Audiology*, *21*(3), 131–141.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *The Journal of the Acoustical Society of America*, *49*(4), Suppl 2:1218+. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4994693

Rhode, W. S. (1978). Some observations on cochlear mechanics, 158–176.

Ruggero, M. A., Robles, L., & Rich, N. C. (1992). Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory-nerve rate suppression. *Journal of Neurophysiology*, *68*(4), 1087–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1432070

Russell, I. ., Richardson, G. ., & Kössl, M. (1989). The responses of cochlear hair cells to tonic displacements of the sensory hair bundle. *Hearing Research*, *43*(1), 55–69. http://doi.org/10.1016/0378-5955(89)90059-2

Russell, I. J., Cody, A. R., & Richardson, G. P. (1986). The responses of inner and outer hair cells in the basal turn of the guinea-pig cochlea and in the mouse cochlea grown in vitro. *Hearing Research*, *22*, 199–216. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3733540

Russell, I. J., & Sellick, P. M. (1983). Low-frequency characteristics of intracellularly recorded receptor potentials in guinea-pig cochlear hair cells. *The Journal of Physiology*, *338*, 179–206.

Russell, I. J., & Sellick, P. M. M. (1978). Intracellular studies of hair cells in the mammalian cochlea. *The Journal of Physiology*, *284*(3), 261–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/731538

Salt, A. N. (2004). Acute endolymphatic hydrops generated by exposure of the ear to nontraumatic low-frequency tones. *Journal of the Association for Research in Otolaryngology : JARO*, *5*(2), 203–14. http://doi.org/10.1007/s10162-003-4032-z

Salt, A. N., Brown, D. J., Hartsock, J. J., & Plontke, S. K. (2009). Displacements of the organ of Corti by gel injections into the cochlear apex. *Hearing Research*, *250*(314), 63–75.

http://doi.org/10.1016/j.heares.2009.02.001.Displacements

Santarelli, R., Starr, A., Michalewski, H. J., & Arslan, E. (2008). Neural and receptor cochlear potentials obtained by transtympanic electrocochleography in auditory neuropathy. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *119*(5), 1028–41. http://doi.org/10.1016/j.clinph.2008.01.018

Scherer, M. P., & Gummer, A. W. (2004). Vibration pattern of the organ of Corti up to 50 kHz: evidence for resonant electromechanical force. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(51), 17652–7. http://doi.org/10.1073/pnas.0408232101

Schmiedt, R. a, Okamura, H. O., Lang, H., & Schulte, B. a. (2001). Ouabain application to the round window of the gerbil cochlea: A model of auditory neuropathy and apoptosis. *Journal of the Association for Research in Otolaryngology*, *3*(3), 223–233. http://doi.org/10.1007/s1016200220017

Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1982). Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *The Journal of the Acoustical Society of America*, *72*(1), 131–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7108035

Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1983). Comparison between the tuning properties of inner hair cells and basilar membrane motion. *Hearing Research*, *10*(1), 93–100. http://doi.org/10.1016/0378-5955(83)90019-9

Sellick, P. M., & Russell, I. J. (1980). The responses of inner hair cells to basilar membrane velocity during low frequency auditory stimulation in the guinea pig cochlea. *Hearing Research*, *2*(3-4), 439–445. http://doi.org/10.1016/0378-5955(80)90080-5

Shaheen, L. A., Valero, M. D., & Liberman, M. C. (2015). Towards a Diagnosis of Cochlear Neuropathy with Envelope Following Responses. *JARO - Journal of the Association for Research in Otolaryngology*, *16*(6), 727–745. http://doi.org/10.1007/s10162-015-0539-3

Shen, H., Hüper, K., & Kleinsteuber, M. (2011). On FastICA algorithms and some generalisations. *Lecture Notes in Electrical Engineering*, *80 LNEE*, 403–432. http://doi.org/10.1007/978-94-007-0602-6_19

Shiraishi, Y., Katayama, N., Karashima, A., & Nakao, M. (2011). Separation of multiunit signals by independent component analysis in complex-valued time-frequency domain. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 4410–4413. http://doi.org/10.1109/IEMBS.2011.6091094

Shore, S. E., & Nuttall, A. L. (1985). High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. *The Journal of the Acoustical Society of America*, *78*(4), 1286–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3840500

Tasaki, I., Davis, H., & Eldredge, D. H. (1954). Exploration of Cochlear Potentials in Guinea Pig with a Microelectrode. *The Journal of the Acoustical Society of America*, *26*(5), 765–773. http://doi.org/10.1121/1.1907415

Teas, D. C., Eldridge, D. H., & Davis, H. (1962). Cochlear responses to acoustic transients: An interpretation of whole-nerve action potentials. *Journal of the Acoustical Society of America*, *34*(9), 1438–1459. http://doi.org/10.1121/1.1918366

Wever, E. G., & Bray, C. W. (1930). The nature of acoustic response: The relation between sound frequency and frequency of impulses in the auditory nerve. *Journal of Experimental Psychology*, *13*(5), 373–387. http://doi.org/10.1037/h0075820

Whitfield, I. C., & Ross, H. F. (1965). Cochlear-microphonic and summating potentials and the outputs of individual hair-cell generators. *The Journal of the Acoustical Society of America*, *38*, 126–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14347603

Wilson, J. P., & Johnstone, J. R. (1975). Basilar membrane and middle-ear vibration in guinea pig measured by capacitive probe. *The Journal of the Acoustical Society of America*, *57*(3), 705–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1123489

Withnell, R. H. (2001). Brief report: the cochlear microphonic as an indication of outer hair cell function. *Ear and Hearing*, *22*(1), 75–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11271978

Yuan, Y., Shi, F., Yin, Y., Tong, M., Lang, H., Polley, D. B., … Edge, A. S. B. (2014). Ouabain-induced cochlear nerve degeneration: synaptic loss and plasticity in a mouse model of auditory neuropathy.

Journal of the Association for Research in Otolaryngology : JARO, *15*(1), 31–43. http://doi.org/10.1007/s10162-013-0419-7

Zheng, X. Y., Ding, D. L., McFadden, S. L., & Henderson, D. (1997). Evidence that inner hair cells are the major source of cochlear summating potentials. *Hearing Research*, *113*(1-2), 76–88. http://doi.org/10.1016/S0378-5955(97)00127-5

Zwislocki, J. J. (1980). Theory of cochlear mechanics. *Hearing Research*, *2*(3-4), 171–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6997254

APPENDIX A: REVIEW OF COCHLEAR MECHANICS

Introduction

This brief literature review begins with an overview of cochlear anatomy relevant to the mechanics that elicit hair cell responses and the consequent neural responses that allow us to hear. The review then spins a brief tale of the history of cochlear mechanics, as a chronological viewpoint gives us insight into the research process that is reflected in the experiments reported in the present study. This process begins with inexplicable data, followed by modeling, and finally physiologic experiments leading to the discovery of amazing biologic mechanisms within the cochlea. The subsequent sections review aspects of cochlear microphonic potential and hair cell transduction mechanisms that are crucial to this study.

Anatomy of the cochlea.

In mammals, sound is funneled through the outer ear and transmitted through the middle ear to the cochlea to stimulate the sensory cells of the ear. The primary roles of the cochlea are to transform mechanical motion into electrical signals and to act as a frequency analyzer. The human cochlea is coiled labyrinth of about 35 mm in length, turned approximately two and a half times forming a 10 mm-diameter structure embedded in the temporal bone at the base of the skull. The cochlea is divided longitudinally into three fluid-filled chambers, called scalae, which vary in size but maintain their spatial relationship along the length of the cochlea from base to apex. From top to bottom, the scalae are named vestibuli, media, and tympani. The scalae vestibuli and tympani contain perilymph –a fluid composition similar to that of cerebrospinal fluid –and are coupled at the very apex of the cochlea via the helicotrema. The scala media is filled with endolymph, similar in composition to intracellular fluid. A special conglomerate of cells that make up the lateral wall of the scala media maintain the ionic concentration of the endolymph and are known as the stria vascularis. The endolymph of the scala media is contained superiorly by a thin, flexible Reissner's membrane and inferiorly by a thicker, more rigid basilar membrane (BM). The function of these membranes is two-fold and relate to the two primary roles of the cochlea, the first being to isolate the endolymph-filled scala media from the perilymph-filled scalae vestibuli and tympani,

maintaining an electrical differential (or voltage) between the chambers crucial to the transformation of mechanical motion into electrical signals. The second role is to stimulate the sensory cells of hearing which sit on the basilar membrane and rely on its motion to be excited.

The input to the cochlea is the piston-like motion of the stapes footplate on the oval window, located at the basal end of the scala vestibuli, which applies pressure to the fluid in this chamber –a pressure which is relieved by an equal and opposite displacement of the round window (RW) located at the basal end of the scala tympani. The difference in pressure at any moment between the scalae vestibuli and tympani is the force that drives the displacement of the BM. The scala media contains an amalgam of cells sitting on the modiolar (or medial) BM, called the organ of Corti (OC) which houses the sensory cells of hearing, the hair cells; peripheral processes of the auditory nerve (AN) which synapse with the hair cells; and the supporting cells. A cross-section of the OC reveals one row of inner hair cells (IHC) and typically three rows of outer hair cells (OHC). Atop each hair cell is a bundle of stereocilia, the base of which is embedded in the connected apical membrane of hair cells –collectively deemed the reticular lamina (RL); and the tips of which are either embedded in the tectorial membrane (TM) or semi-freely moving in the endolymphatic fluid in the space between the reticular lamina and tectorial membrane. The tectorial membrane is a gelatinous flap anchored only on its modiolar edge to the limbus, joining it to the reticular lamina. Therefore, the transverse wave-like displacement of the BM also causes a slightly-delayed displacement of the tectorial membrane. The out-of-phase oscillation of these two membranes results in a shearing motion, displacing the stereocilia of the hair cells in a medial-lateral direction, opening ion channels residing in the membranes of the cilia, and allowing an influx of charged ions from the endolymph into the body of the hair cells. Current produced by this ionic flow generates a voltage across the cell wall membrane, enough of which will trigger the release of neurotransmitters to the synapsed AN. If enough neurotransmitter is received by the auditory nerve in a given time, an action potential (AP) is generated and propagated to the cell body residing in the spiral ganglion, and then sent through the middle of the cochlear spiral (the modiolus), down the length of the neuron to the central auditory system, triggering a series of neural signals that ultimately reach the auditory cortex.

There are about 15,000 hair cells in humans, with 3 times as many OHCs as IHCs. The role of IHCs and OHCs are incredibly different in the mammalian cochlea. Inner hair cells are responsible for the majority of afferent signaling with 90-95% of afferent AN fibers connecting to IHCs. The afferent fibers innervating IHCs are thick and myelinated and called Type I fibers. There are typically 20-30 Type I fibers synapsed on one IHC. The remaining 5-10% of afferent AN fibers are thin, unmyelinated, and take a long, winding basal pathway to innervate multiple OHCs; the function of these Type II fibers is still debated. Additionally, both hair cell types receive efferent innervation from the central auditory system, though OHCs receive the majority of efferent fibers. Efferent activity allows the OHCs to fulfill a role in fine-tuning the sensitivity of the cochlea. Another important role of the OHC is cochlear amplification. Low voltage changes from low input levels cause the OHCs to expand and contract due to a specialized protein, Prestin, found in the cell wall membrane; enhancing the displacement of the BM and tectorial membranes, effectively increasing the input to IHCs.

History of Cochlear Mechanics.

While studies of the anatomy of the cochlea go back to the ancient Greeks, modern studies of its function were pioneered by Georg Von Bekesy in the 1940s. In experiments on human and animal cadavers, Bekesy was able to describe the motion of the cochlear partition in response to sound. His experimental approach included dissection of the temporal bone *postmortem*, replacing the oval and round windows with rubber, and attaching a vibrator to one of the windows. Bekesy opened the top of the cochlea and sprinkled silver particles onto Reissner's membrane, allowing him to view its motion in response to vibration. What Bekesy saw was what we now refer to as the travelling wave. To a pure-tone stimulus, or fixed-frequency vibration, he found that the cochlear partition oscillated to the same frequency as the stimulus, but the amplitude of the wave displacement varied across the length of the partition: the wave grew slowly from the base, peaked, and then rapidly decayed toward the apex. When the frequency of the stimulus varied, he found that the location of the peak amplitude changed while the general pattern remained: the higher the stimulus frequency, the more basal the peak location and low frequency tones show a very gradual sloping increase in displacement from the base to a peak near the

apex. In additional experiments, Bekesy observed BM displacement at one point along the partition as he varied stimulus frequency. Again showing that the basal cochlea responded to higher frequencies and vice versa, but this experiment also showed a much steeper slope on the higher-frequency side of the response displacement curve and a very shallow slope on lower-frequency side. Bekesy inferred that the BM acts as a low-pass filter with poor frequency selectivity, due to the shallowness of the low-frequency tails of the BM tuning curves. Bekesy's remarkable findings are attributed to the beginnings of the field of cochlear mechanics.

In 1965, Kiang et al. studied the tuning curves of the AN in cats and discovered that peripheral auditory neurons had much sharper tuning curves than the BM tuning curves put forth by Bekesy. Then in 1967, Johnstone & Boyle applied the Mossbauer technique to measure BM vibration. The Mossbauer technique was more sensitive than Bekesy's light microscopy approach and it allowed them to study on live animals. Johnstone and Boyle discovered that the slope of the BM transfer function in the guinea pig was much steeper than that found by Bekesy, at least near the base where they measured. Since then, it has been shown that in order to get viable BM measurements, the subject must be alive and the cochlea in good shape (Sellick et al., 1982). The original discrepancy between Bekesy's BM tuning and the AN was not quite reconciled by Johnstone and Boyle, but their findings spurred William Rhode to further experiment on the BM using the Mossbauer technique. Using squirrel monkeys, Rhode confirmed the transfer function slopes published by Johnstone and Boyle but in the mid-frequency regions as well. He was able to stimulate the BM using lower levels than those required by Bekesy's light microscopy, and found that at the lowest levels he could stimulate (70-80 dB SPL) the transfer function saturated at characteristic frequencies, showing a compressive nonlinearity not previously found in any study. A number of follow-up studies to Rhode's publication in 1971 argued against this nonlinearity (Johnstone et al., 1970; Kohlloeffel, 1972; Helfenstein, 1973; Johnstone & Yates, 1974; Wilson & Johnstone, 1975), however, Rhode argued that a nonlinear BM was "fairly consistent [with] of nonlinear phenomena observed neurally, psychoacoustically, and mechanically" (Rhode, 1978).

At the same time Bekesy was conducting his experiments, an English physicist, who was known for his work in astronomy and had just begun research on the ear, proposed a ridiculous idea to the hearing science community; that there was an active mechanism in the cochlea. He reasoned that a passive system could never produce the sharp tuning seen at the level of the AN or the human ability to discriminate between small frequencies, and that this amplification must occur before the AN (Gold, 1989). Thomas Gold's proposal was immediately shunned by the community and ignored until Kemp provided evidence of a positive feedback system in 1978, when he formally introduced objective measures of tinnitus and otoacoustic emissions. The idea that this mechanism was an active biological process was also evidenced by the fact that this nonlinearity is only seen in live preparations. In 1984, William Brownell recorded an OHC under a microscope and evoked a motile response in an outer hair cell which led to the idea that OHCs are the cochlear amplifier. And thus, an entire area of study within the realm of hearing science was born to argue the source of cochlear amplification. Fortunately, the stimuli and responses used in the present study likely did not evoke the cochlear amplifier so this subject will not be reviewed further.

The Cochlear Microphonic

In 1930, Wever and Bray discovered an electrophysiologic response to acoustic stimuli recorded from the AN**.** They described a potential which alternated polarity at the same frequency as the stimulus with such accuracy that when they stimulated the ear with speech and recorded the response, the electrode recording could be played back and the speech understandable. They found that this occurred only in healthy cochlea and concluded that this phenomenon was neural in origin. In 1931, Adrian followed up on this experiment and concluded that this response could be recorded from anywhere around the cochlea, not just the AN, and argued that the response is likely but may not, in fact, be neural in origin. Hallpike & Rawdon-Smith recorded the Wever-Bray phenomenon in cats with electrodes placed at various locations along the cochlear partition and found that the response amplitude was greatest in the base for high frequencies and greatest in the apex for low frequencies, though in their 1934 paper argued that the response could be either neural or non-neural. Davis and colleagues, on the other hand, proposed that the

response origins were not neural, but rather the sensory cells in the organ of Corti, and they argued that this could explain the results of Wever and Bray just as well. Davis's model would prove correct and the response known as the cochlear microphonic.

The primary function of the cochlea is to transform mechanical movement into electrical signals. The link between the mechanical motion of the BM and the electrical response in the AN is the hair cell. The function of the hair cell is often spoke of in terms of resistance modulation as put forth by Davis in his battery theory of the cochlea 1958. Davis described a model showing a constant flow of current outward from the scala media though the hair cells. When the resistance is constant, i.e. when ion channels are closed, electrical current that flows through the hair cells exits through the lateral membranes of the hair cells back into the scala media and the endocochlear potential is maintained. When the resistance of the hair cell membrane is reduced, i.e. when ion channels are opened, the endocochlear potential and negative hair cell intracellular resting potential form a gradient across the apical membrane of the hair cells (or reticular lamina) such that current increases and flows toward the basilar membrane, causing changes in the endocochlear potential. Animal experiments were able to confirm aspects of the model, such as the resistance values of the hair cell membranes and current flow directions (Honrubia & Ward, 1970; Johnstone et al., 1970); or that the AN fires when the CM is negative (Kiang, 1965; Teas, Eldridge, & Davis, 1962). In the 1970's, Dallos and colleagues compared CM recordings from inner and outer hair cells and determined that OHCs are the dominating contributor to the CM (Dallos et al., 1972; Dallos 1973; Dallos & Cheatham 1976). Since then, the CM has been typically referred to in the literature as current flow through OHCs alone, though it truly reflects receptor current though both IHC and OHC.

The CM has typically been recorded in animals from an electrode placed in the RW niche. The CM is an extracellular potential and is based on the notion that mechanical stimulation due to the introduction of an acoustic stimulus modulates the transducer current of the hair cell which effect changes in the extracellular current flow. This produces voltage changes across the scalae membranes measured by the electrode (Davis et al., 1958; Hudspeth & Corey, 1977). The CM is also a gross potential, meaning the response recorded by the electrode is a linear combination of responses from any and all stimulated

cells. Due to principles of electric fields, each response is a vector weighted by its distance to the electrode and impedances of the media it passes through (e.g. cochlear membranes will have a different impedance than bone which will affect the response vector from a given hair cell location). The literature has shown that the CM, even at lower stimulus frequencies, is dominated by OHCs at the base of cochlear partition due to their proximity to the electrode. Patuzzi et al. (1989) ablated the 3rd and 4th turns of the guinea pig cochlea and found little change in the CM to frequencies above 300 Hz when recorded from the RW. Damage to hair cells in the base of the cochlea with streptomycin and noise exposure resulted in dramatic reductions to the RW-recorded CM. To low-frequency stimuli, it is thought that the majority of the BM at the base of the cochlea is moving in-phase (He et al., 2012). This results in stimulation of basal hair cells even to low frequencies. Thus, it is the low-frequency tails of the tuning curves of basal OHCs that are thought to be dominating the response. If this is the case, then the CM is a passive response that does not evoke OHC motility and amplification. Indeed, in a study by Cheatham et al., (2011), the CM in Prestin knock-out mice was not significantly different than that of mice with healthy OHCs.

While it seems that the cochlear microphonic could be used in some way to assess the health of OHCs, the interpretation of the response is not yet fully understood and the technical challenges in recording the response, such as the difficulty in recording responses from apical hair cells, has hindered development of the CM into a diagnostic auditory-evoked potential. Withnell (2001) pointed out several issues with using the CM in humans. The first is that most auditory-evoked potentials in the clinical, such as the ABR and electrocochleography, tend to use stimulus frequencies in the range of 0.5-4 kHz. Most recordings of the CM are with lower frequency stimuli (<1 kHz) and even a 200 Hz tone results in a response that is dominated up to 98% by OHCs with a characteristic frequency of 8 kHz or higher (Patuzzi et al., 1989). Withnell stated that solving this problem by placing an electrode proximal to the characteristic frequency of a target region of hair cells, (i.e. not a RW recording), still poses the problem of phase-cancelations of a response from the targeted area. The BM will have the greatest displacement and consequently large phase rotations at the characteristic frequency site of the stimulus. If an electrode is placed near the stimulus frequency location on the cochlear partition, the response from this area may

be canceled due to summation of out-of-phase neighboring hair cells. This poses a dilemma that would have to be solved in order for the CM to be used to assess the health of OHCs.

In 1992, Ponton and colleagues proposed the use of a high-pass noise to 'mask' the CM and obtain location-specific responses. While masking has been used to target location-specific responses in other auditory-evoked potentials such as the CAP (Teas et al., 1962; Eggermont, 1976; Elberling, 1974; Naunton & Zerlin, 1976; Spoor et al., 1976; Evans & Elberling, 1982; Shore & Nuttall, 1985; Teas et al., 1962; Earl & Chertoff, 2012) and ABR (Conijn et al., 1992; Boettcher et al., 1995; Klein, 1986; Oates & Stapells, 1997), the term 'masking' has traditionally been used to refer only to the suppression of neural responses. However, OHC responses can also be suppressed, though the mechanism is different than that of the auditory nerve (Ruggero et al., 1992). It is thought that a suppressing sound near the probe tone frequency will change the operating point of the transducer curve of the hair cell by keeping the stereocilia constantly deflected such that any addition deflection due to stimulus tone will be one-directional and small relative to the response from unsuppressed hair cells. Some argue that the term masking cannot be used to describe suppression of hair cells, but in the present study we refer to this suppression as masking, as did Ponton. In 2012, Chertoff and colleagues developed a high-pass masking noise paradigm and showed evidence that responses from the base can be masked in order to obtain CM responses from farther down the cochlear partition in the gerbil. Chertoff et al. (2014) and Kamerer et al. (2016) successfully used this masking paradigm in gerbils to record the CM at 762 Hz and 45 & 85 Hz tones, respectively. In 2017, Charaziak and colleagues used suppressing tones either slightly above or below the stimulus frequency and subtracted these responses from the CM without suppressing tones to obtain location-specific CM responses in chinchillas. These studies have provided hope for the future use of the CM as a tool to assess the health of OHCs at specific locations along the cochlear partition.

The CM has been recorded in humans even if it is not yet used clinically. It was first recorded in humans by Perlman in 1941. Because the CM is small, the more proximal the electrode is to the source, the better the recording. Therefore, the best way to record the response in humans is transtympanically with a needle electrode placed on the promontory (Eggermont & Odenthal 1974; Santarelli et al., 2008).

In fact, the CM can be recorded using clinical tools that already exist. Electrocochleography (ECochG) is a method of recording which optimizes wave I of the ABR: the CAP of the auditory nerve. Using this technique, the CM and the summating potential (SP) can also be recorded. The SP is another hair cell receptor potential that is reflected in the DC shift of the ECochG response. In response to tonal stimuli, hair cells produce an AC response that reflect the displacement pattern of the BM; this is the CM. They also produce a DC response which follows the stimulus envelope (Tasaki, Davis, & Eldredge, 1954). Since this is a DC product of an AC stimulus, it is inherently a distortion product due to nonlinearities in the hair cells, and thought to be a distortion-product of the CM response itself (Whitfield & Ross, 1965; Johnstone, Johnstone, & Pugsley, 1966; Davis, 1968; Engebretson & Eldredge, 1968; Honrubia & Ward, 1969). Nonlinearities such as the SP are due to the transduction mechanisms of the hair cells and are important to the present study, which uses distortion-products to differentiate cellular sources of the low-frequency CM.

Hair Cell Transduction

The stereocilia of the OHCs are embedded in the TM. Movement of the BM upward toward the scala vestibuli causes a pivoting of the reticular lamina toward the modiolus. As the TM moves upward as well, there is very little radial movement of the TM resulting in a shearing pattern of these two membranes deflecting the hair cell stereocilia toward the stria vascularis. The deflection of the stereocilia in the lateral (or excitatory) direction mechanically pulls on proteins that connect individual stereocilia, known as tip-links. Stretching of the tip-links is thought to open ion channels in the membranes of the cilia at the site of the links, allowing an influx of K+ ions into the cell body. Potassium ions follow the potential gradient from the endolymph of the scala media to the relatively negative intracellular fluid of the hair cell. The build-up of K+ in the cell body generates a receptor potential, reflected indirectly by the CM and SP. Enough receptor potential depolarizes the cell resulting in a release of neurotransmitters from the basal membrane of the hair cell into the synaptic cleft. Dendrites on the other side of the synaptic junction receive the neurotransmitter generate a potential, exciting the innervating AN fibers. Conversely, downward movement of the BM causes the stereocilia to be deflected toward the modiolus, which

releases tension in the tip-links, closing the ion channels. The efflux of K+ ions and influx of Ca++ (calcium channels are thought to open at the onset of stimulation) hyperpolarizes the cell. It is important to note that, since OHC stereocilia are mechanically coupled to the TM, their CM response is in-phase with BM displacement.

OHC stereocilia are coupled to the TM, therefore movement of the TM in the radial direction will directly and mechanically open and close of ion channels in the cilia membrane. Inner hair cell stereocilia, on the other hand, are not coupled to the TM, therefore it is of some debate how exactly IHCs are excited. Guinan (2012) proposed two mechanisms which could excite IHCs depending on the relationship between characteristic frequencies of the cells and the stimulus frequencies and levels used to excite them. The shearing motion between TM and RL is just one of said mechanisms and how this shearing affects the IHC stereocilia is dependent on stimulus frequency. A simplistic answer is that the shearing motion of the RL and TM causes viscous drag in the fluid of the space between them producing enough force to mechanically deflect the stereocilia of the IHCs. At low-frequencies, the IHC stereocilia are sensitive to the velocity of TM relative to the RL, meaning the IHC response is 90° out-of-phase with the OHC response (Sellick & Russell, 1980; Nuttall et al., 1981; Russell & Sellick, 1983; Russell et al., 1986; Patuzzi & Yates, 1987; Freeman & Weiss, 1990a, 1990b; Cheatham & Dallos, 1999; Jia et al., 2007). Conversely, many of the same studies found that the IHC response for high-frequency stimuli (>300 Hz) was in-phase with the OHC response, or sensitive to BM displacement. This is due to the mechanical properties of the TM. Cochlear models consider the TM as a mass attached to springs at the modiolus and OHC stereocilia. At low frequencies, the TM movement is dominated by its elasticity and at high frequencies, its mass. At low-frequencies, the TM has been shown to move very little, radially, remaining in-phase with the BM (Chan & Hudspeth, 2005; Karavitaki & Mountain, 2007b). Models of the TM show that it resonates at a frequency lower than the local characteristic frequency, such that at high frequencies, the TM motion is out-of-phase with BM and predicted phase relationships would not be the same as that produced by the typical shearing motion (Allen, 1980; Zwislocki, 1980). The second proposed excitatory drive for the IHC occurs at low stimulus levels at the location where hair cells'

characteristic frequency is that of the stimulus frequency–where cochlear amplification is evoked (Guinan, 2012). Currently the strongest evidence for the origin of cochlear amplification is the somatic motility of OHCs. If this is the case, then OHC motility may drive excitation for IHCs as well as the shearing drive. If it is assumed that the RL is a rigid bar that pivots about the head of the pillar cells – supporting cells in the organ of Corti separating the OHCs and IHCs –then OHC motion would change the RL-TM gap (Scherer & Gummer, 2004; Nowotny & Gummer, 2006, 2011). Chiardia and colleagues (2009) modeled how expansion and contraction of the OHCs would squeeze and expand the gap between the RL and TM, causing changes in the fluid flow and deflecting IHC stereocilia. How this would affect IHC responses is complicated, since the shearing motion would excite IHCs and induce OHC motility which in turn would reduce the effect of the shearing motion on IHCs. This is likely not an excitatory drive for the hair cells excited in the present study, as the stimuli that were used were 80 & 90 dB SPL: much higher than would evoke OHC motility. Guinan (2012) also proposes two mechanisms by which IHCs are inhibited, which affect the cell's hyperpolarized state, both of which are related to the coupling of OHC stereocilia with the TM. The first theory is simply that the stereocilia are not perpendicular to the TM at rest. If the stereocilia are embedded at an angle, then deflection of the cilia in one direction could significantly change the size of the RL-TM gap, altering fluid flow through the gap and affecting IHC stereocilia. The second theorizes that when the BM moves upward, the OHC stereocilia push into the gelatinous exterior of the TM without moving the TM as much as if it were a rigid structure, resulting in a squeezing of the RL-TM gap pushing fluid toward the modiolus and inhibiting IHCs. All these mechanisms result in interesting phase relationships between IHCs and OHCs and are important to the results of the present study. Additionally, these mechanisms may explain in part the asymmetric relationship between depolarizing and hyperpolarizing states of the IHC which have been found in intracellular studies of IHC transduction.

In the 1970s, the advancement of neuroscience technology in the field of hearing allowed hair cells receptor potentials to be studied inside the cell, instead of indirectly via gross potentials like the CM. Inner hair cell responses were first recorded by Russell & Sellick in 1978 using a microelectrode which

could penetrate the cell wall of a single hair cell *in vivo*. Intracellular recordings in live animals are incredibly tedious and the fidelity of the responses rely heavily on maintaining the integrity and health of the cochlear structures. Therefore, only a couple of research groups have these measurements: Russell & Sellick and Dallos and colleagues, with the majority of experiments done on the guinea pig. The guinea pig model is optimal for such recordings because the orientation of the cochlea in the auditory bulla exposes the apical turns, as opposed to gerbils, where only the RW and base of the cochlea are easily-accessed. Quite of number of *in vitro* recordings followed these groups of *in vivo* recordings (Sellick & Russell, 1980; Nuttall et al., 1981; R. Patuzzi & Sellick, 1983; Sellick, Patuzzi, & Johnstone, 1983; Patuzzi & Yates, 1987; Dallos & Cheatham, 1989; Russell, Richardson, & Kössl, 1989; Jia et al., 2007). Most of their findings mirrored that of *in vivo* experiments, however, many processes of hair cells rely on their health, hence the focus of this review will be on *in vivo* results.

When hair cells are stimulated with acoustic tones, the hair cells generate a receptor potential which is reflected in a phasic, frequency-following AC response. This time-domain waveform is distorted; it contains higher order harmonics of the fundamental (or stimulus) frequency and a DC component (Dallos & Cheatham, 1992). This distortion arises from nonlinearities in the transduction mechanisms and are their relative size is dependent on the relationship between the characteristic frequency of cell and the stimulus frequency and level. The current flow through the cells' basolateral membrane is regulated by the ion channels in the stereocilia. This produces the receptor potential. The conductance of the basolateral membrane modifies the receptor potential due to the capacitance of the membrane. The membrane acts as a low-pass filter, because of the time-constant of the membrane's ability to expel ions. Thus, AC component of the potential is attenuated at frequencies above this time-constant, which sits around 1200 Hz (Russell & Sellick, 1978; Palmer & Russell, 1986). At higher frequencies, the DC is the only remaining component. The intracellular AC amplitude is comparable between the two hair cell types, which explains why extracellular potentials like the CM are dominated by OHC (as there are three times as many OHC than IHC). A hair cell's response amplitude to a given stimulus frequency, however, is different between the two types of hair cells. The low-frequency slope of

the AC tuning curve for IHC is 6 dB steeper than that of OHCs. Additionally, at low signal levels, IHCs only display a positive DC amplitude, or strong depolarizing phase, while OHCs show stronger hyperpolarization (negative DC potential) to stimulus frequencies below the characteristic frequency which reverses to a stronger depolarizing (positive DC potential) with increasing frequency (Dallos, 1985a). At higher signal levels, both hair cells display only a positive DC response regardless of frequency.

The level-dependence of the hair cell response is usually shown as a transducer function; a plot of peak magnitudes as a function of condensation and rarefaction sound pressure levels (see Fig. 14). The transducer function, or curve, also exposes any asymmetry between the hyperpolarizing and depolarizing states and thus will predict the DC responsiveness. At low sound pressure levels, the output of the hair cell response grows linearly and eventually saturates. In some models, this saturation is not a complete halt to the growth of the response amplitude, but more of an extreme compression (Cheatham & Dallos, 1994). Other models have shown a saturation point followed by a small decline in the curve (Dallos 1985a). Responses from the linear portion of the curve will not produce higher-order harmonic distortion; it is in the compressive phase of the curve where harmonic distortion is produced. It is also important to note that high-frequency AC responses are still filtered out after they go through the transduction process, therefore the AC response predicted by the transducer curve for frequencies above ~1.2 kHz will not match the AC recorded in the CM. As shown in Fig. 14 different research groups have obtained slightly different transducer curves for inner and outer hair cells. This may partly be due to differences in recording methods, however, it is more likely due to 1) the relationships between characteristic frequency of the cell and the stimulus frequency used to elicit the response and 2) the health of the cochlear structures, as it has been shown that the operating point of the cell may change due to the health of cochlea which would affect this curve. The resting potential of the cell, or the output (V) of the cell in the absence of stimulation, is the operating point, which can be affected by changes to the endolymphatic fluid and endolymphatic hydrops (Serjani, Salt, Gill, Hale, 2003). Changes to the operating point affect at what levels the output growth is linear or compressed. The methods used to record intracellular responses

require drilling a hole through the bone surrounding the labyrinthine structures within the cochlea, and puncturing the scala membranes with a needle electrode, which has been shown to produce endolymphatic hydrops. Therefore, it is difficult to conclude the exact reason for the differences between these recordings. The asymmetry and saturation points of a given hair cell transducer curve will predict whether the DC component from that cell will be positive or negative, and vital to the present study: what higher order harmonics each cell will produce. A symmetric, saturating function such as that of Fig. 14(c) will produce odd harmonics (F3, F5, etc.) while an asymmetric, saturating function like Fig. 14(i) will produce all harmonics (F2, F3, F4, etc.).

The transduction mechanisms of both IHCs and OHCs are important to the generation of the CM. The CM is an extracellular potential that reflects voltage changes across scala membranes. This voltage is due to changes in current flow through hair cells as they modulate their receptor potentials. The CM that is recorded from the RW is a gross summation of all the voltage changes in the cochlea weighted by their distance to the electrode. Distortions in the CM reflect the transduction mechanisms of the hair cells that generate the potential. Features like the sign of the DC component, and specific distortion frequencies may give us clues as to which of these hair cells are the source of the CM being recorded, and maybe the characteristic frequencies, or locations, of the sources. I hope this brief review clarified some of the physiologic mechanisms behind the reasoning used to draw conclusions from the results of the present study.

APPENDIX B: INDEPENDENT COMPONENT ANALYSIS

Introduction & the Blind Source Separation Problem

This section provides a basic explanation of an analysis used in Experiment 4. I hope the reader will use this information to 1) appreciate the innovation and broad applications of ICA and why we chose this approach, 2) gain a slightly better understanding of the analysis itself, and 3) realize its complexity and why a non-expert in linear algebra or the signal processing field may not have gleaned the results he/she wanted from the analysis.

In the last decade or so, there has been a boom in the development of tools that use linear algebra to perform signal processing. Methods that solve the problem of blind source separation are innovating technologic and engineering fields, and are also making their mark in healthcare advancement. So, what is the blind source separation problem? In short it is separating source signals from a set of signal mixtures without (or with very little) prior information about the sources. Imagine two people in a room talking at the same time. And imagine you have two microphones set up at different locations in the room, recording the two speech signals. The output of each microphone will be a mixture of the two speech signals, weighted spectrally and temporally by each speaker's distance to the microphone. We can write an equation for each of these mixtures as the following:

$$\begin{aligned} x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) \\ x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) \end{aligned} \tag{8}$$

where x_1 and x_2 are the two mixtures, whose value at time (t) is the linear combination of the two source signals, s_1 and s_2, weighted by some coefficient, a_{ij}. The sources and the values of the coefficients are unknown. All you have are x_1 and x_2 and you know there are two sources. What if you want to know the original sources? This is the problem of blind source separation.

Independent component analysis (ICA) has become a standard method of solving the blind source separation problem (Shen, Hüper, & Kleinsteuber, 2011). Since its birth in the 1990s (Cardoso & Souloumiac, 1993; Comon, 1994), the original ICA algorithm has been tweaked for efficiency and modified for a variety of signal processing situations. These methods may use different features of the

signals, such as kurtosis or entropy; or different methods of estimation, such as gradient ascent or infomax, to find the sources signals within the mixtures. Regardless of variation, all ICA algorithms are distinguished from other methods of blind source separation by two assumptions about the source signals; they must be *statistically independent* and *nongaussian.*

Assumptions of ICA

Statistical independence means one event cannot predict another. In this case of ICA, this means the values of one source signal does not convey any information about the values of another. In reality, this assumes that the two source signals come from different physical causes. Nongaussianity refers to the distribution of the source signals. ICA works by maximizing the non-normality of the source signals and relies on principles of the central limit theorem stating the distribution of the sum of the two sources will be more Gaussian than either of the independent source distributions alone. If both of the sources have a Gaussian distribution, their sum distribution may not be different enough from either source distribution for ICA to find the necessary information to separate the two sources. Additionally, for the central limit theorem to hold true to the data, the sources must be additive, therefore ICA assumes that the sources are linearly mixed (e.g. Eq. 8). ICA sees each mixture as a random vector in a matrix. It uses some metric of nongaussianity depending on the specific algorithm, to weight an unmixing matrix that pulls source vectors out of mixture matrix. Thus, another important assumption of ICA is that the number of sources equals the number of mixtures. If the number of mixtures is greater than the number of sources, the case is overdetermined and ICA may perform poorly. In data sets for which the number of mixtures is less than the number of sources –an underdetermined model –traditional ICA will not perform at all. There exist variations of ICA that specifically deal with situations of over- and under-determined data sets, however, these were not explored in this study.

Preprocessing

Before analysis, data used for ICA are typically preprocessed. The reasons for preprocessing are to simplify the algorithm, reduce the dimensionality of the problem, reduce the number of covariates to be estimated, and highlight features of the data or that cannot be explained by the mean and covariance (Li

& Zhang, 1998). The primary methods of preprocessing are centering and whitening. Centering eliminates the mean information simply by subtracting it from the data. Geometrically, this simply translates the center point to the coordinate origin. Whitening is the process of removing any correlations in the data, or forcing the two mixtures to be uncorrelated, which makes the ICA process faster and easier. Figure 23(a) is a plot of the values of two example source signals against each other at each moment in time such that s_1 is the abscissa and s_2 is the ordinate of the coordinate of each data point. Note that the distribution is uniform and square, or orthogonal, because the two sources are statistically independent and nongaussian. Figure 23(b) is a plot of two arbitrary linear mixtures of s_1 and s_2. Note the rhombus-shape to the data, implying some correlation between the two mixtures. If the mixtures are whitened and plotted again, we get Fig. 23(c). The distribution of the data is now orthogonal and all ICA has to do to find the two sources is "rotate" this representation some solvable amount. Geometrically, whitening is a just a scaling transformation of the data to an orthogonal state, which removes a number of covariates with which ICA has to work.

Theory

In the example of Eq. 8, there are four different vectors of *a* coefficients which weight each of the two sources in each of the two mixtures as a function of time and thus make up the mixing matrix, **A**. The inverse of this matrix, then, is an unmixing matrix (**W**) which can be used to solve for the source signals:

$$\begin{aligned} s_1(t) &= w_{11}x_1(t) + w_{12}x_2(t) \\ s_2(t) &= w_{21}x_1(t) + w_{22}x_2(t) \end{aligned} \tag{9}$$

where w_{ij} is the weighting coefficient for each mixture at each point in time to solve for a given source signal. The goal, therefore, of ICA is to estimate the matrix **W** so that it weights the mixtures in such a way that the estimated sources are as independent as possible. Graphically, **W** rotates the data in Fig. 23(c) so that it matches the data in Fig. 23(a).

The next question is how does ICA estimate **W**? Independent component analysis is trying to solve for source components that are as independent as possible, and according to the central limit theorem, independent components are maximally nongaussian components. Therefore, if ICA can find

the maximum nongaussianity within a mixture, it has found an independent source component. How ICA finds maximum nongaussianity is where the analysis diverges into numerous different types and becomes quite complicated.

FastICA

In order to find this maximum nongaussianity, ICA relies on higher-order features of the data, like kurtosis and nonlinear correlations, and cannot use information in the first two moments of the distribution, the mean and covariance. Different types of ICA use different pieces of information. Additionally, different numerical methods are used to compute the optimization algorithm. Put forth by Aapo Hyvärinen in 1999, FastICA is probably the most popular ICA algorithm, liked for both its accuracy and speed.

FastICA relies on the minimization of mutual information to find independent components. Mutual information is a measure of the independence of random variables. The matrix, **W,** is generated such that the mutual information in the resulting estimates sources is minimized. A quantifiable measure of the independence of variables is entropy –a measure of how uniform the distribution is of a set of values. Maximum entropy occurs when there is complete uniformity. The joint entropy of a set of variables depends on the uniformity of the values in the set, and is measured from zero to unity. A fair example of unity as defined by the joint entropy of two variables is Fig. 23(a). A way of obtaining mutually independent signals –or minimizing mutual information –is to maximize entropy. Comon (1994) and Hyvarinen (1999) were able to use this concept in ICA by differentiating the equation for finding the entropy of a random vector with a given probability density function. The normalization of the differentiated entropy gives rise to negentropy –a measure of nongaussianity. Thus, the matrix, **W**, is now determined such that negentropy is maximized.

ICA next has to estimate negentropy for a given set of variables. There are several ways to approximate negentropy but FastICA uses the maximum entropy principle (Hyvarinen 1999). This approximation relies on the cumulative distribution function, *g*. The cumulative distribution function is the area under the probably density function (or distribution). A signal can be transformed by its own

cumulative distribution function which will result in a uniform distribution, or maximum entropy. In order to find one independent component in a mixture, ICA maximizes entropy and works backward to find the signal who has that distribution function. If we want to find more than one independent component we will have more than one *g* function (which is usually the case) and thus, some kind of optimization algorithm must be introduced to the problem.

Many ICA methods use gradient ascent, which incrementally changes the unmixing coefficient values in a direction that maximizes whatever measure is being used (kurtosis, entropy, etc.). Since it is doing this for several variables at the same time, it is trying to move in different directions simultaneously, so it follows a gradient of steepest change until it reaches convergence. Hyvarinen argues that a fixed-point iteration scheme is faster and more user-friendly than methods that use gradient ascent. In a fixed-point algorithm, computations are made in large data blocks and run in parallel and converge faster. In gradient ascent methods, the cumulative distribution function, *g*, must be known *a priori.* In FastICA, which uses the fixed-point scheme, an estimate of *g* is not necessary to find independent components, but can be used by the algorithm, if known, to speed up convergence and increase accuracy.

Concluding remarks

ICA has been used extensively in biomedical signal processing, including EEG, where it can be used to filter out noisy signals like eye blinks from the data (Lindsen & Bhattacharya, 2010; Shiraishi et al., 2011). ICA has been used to localize epileptic activity for pre-surgical evaluation in both time-domain data from EEG and spatial data from fMRI (Hunyadi et al., 2013). It has even been used in fetal magnetocardiography, where heartbeat signals from mother and baby need to be separated in order to assess fetal health (James & Lowe, 2003; de Araujo et al., 2005). In the field of hearing science, ICA has been used to improve signal processing in cochlear implants (Kokkinakis & Loizou, 2008), but has never been used to process the electrical response from the cochlea. In theory, ICA will be 100% accurate if the assumptions of the data are met, however, this is much more difficult in practice. Since we don't know what the sources are, how do we know if they are truly independent? Different ICA algorithms have been created to deal with a variety of signal and signal-mixing situations, but none are perfect for real-life data.

Different algorithms can lead to different results using the same data set, and since the problem of blind source separation is that we don't know *a priori* what the sources are (which is why we need ICA), it is hard to know which algorithm is best.

APPENDIX C: FIGURES AND TABLES

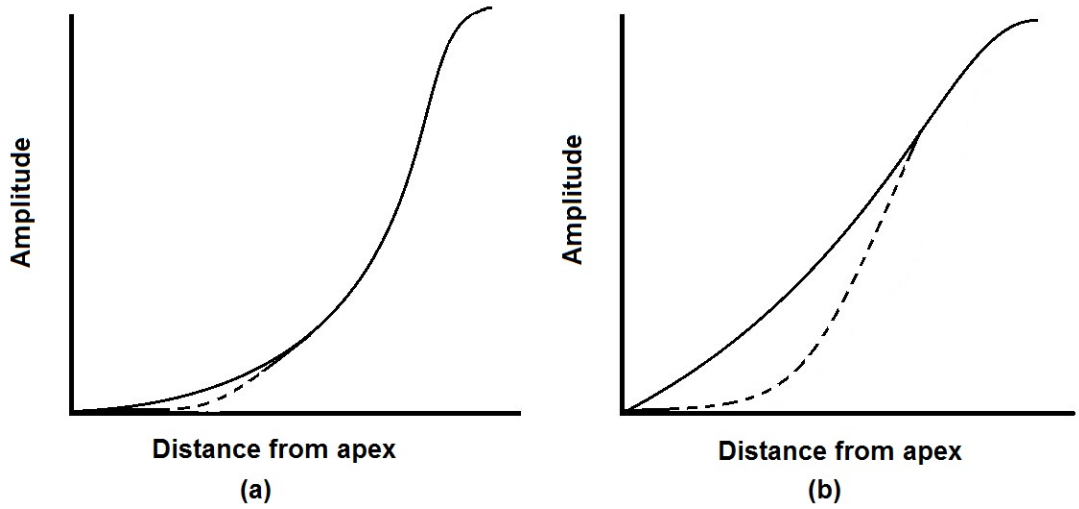

Figure 1. Schematic of a Cumulative Amplitude Function (CAF) seen at 762 Hz (a). The solid line predicts the CAF of a healthy ear while the dotted line predicts the CAF of an ear with damage to outer hair cells in the apex of the cochlea. Fig 1(b) is the hypothesized low-frequency CAF for a healthy ear (solid line) and apically-damaged ear. A more linear CAF will show damage to the apex more clearly than the typical sigmoidal CAF.

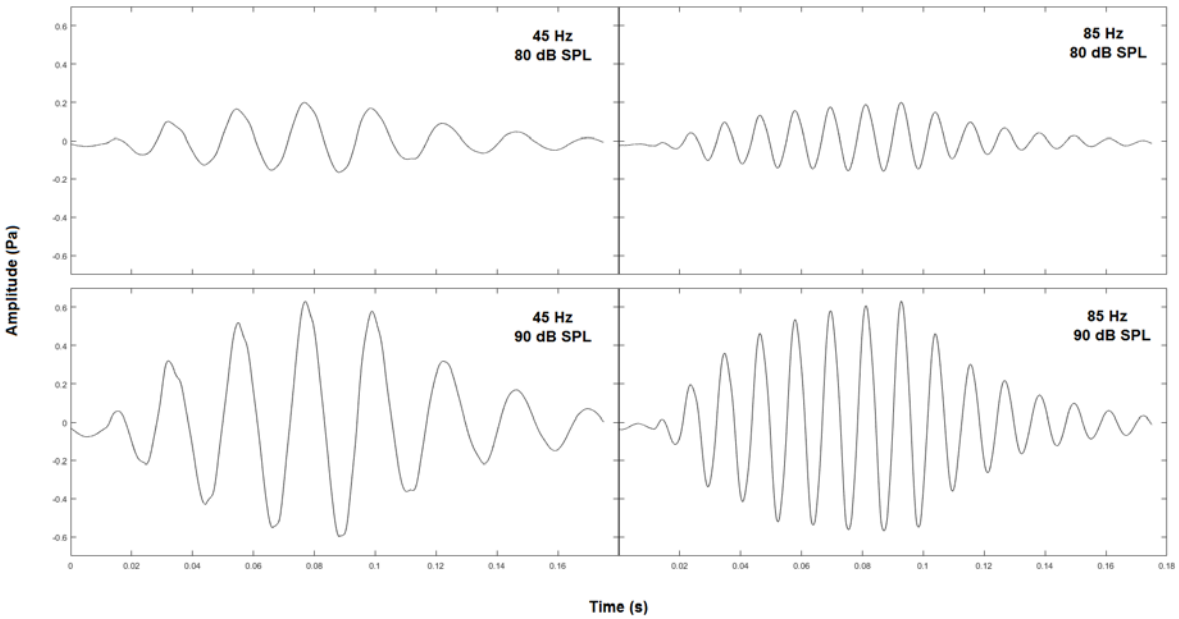

Figure 2. Low-frequency stimuli in the time-domain.

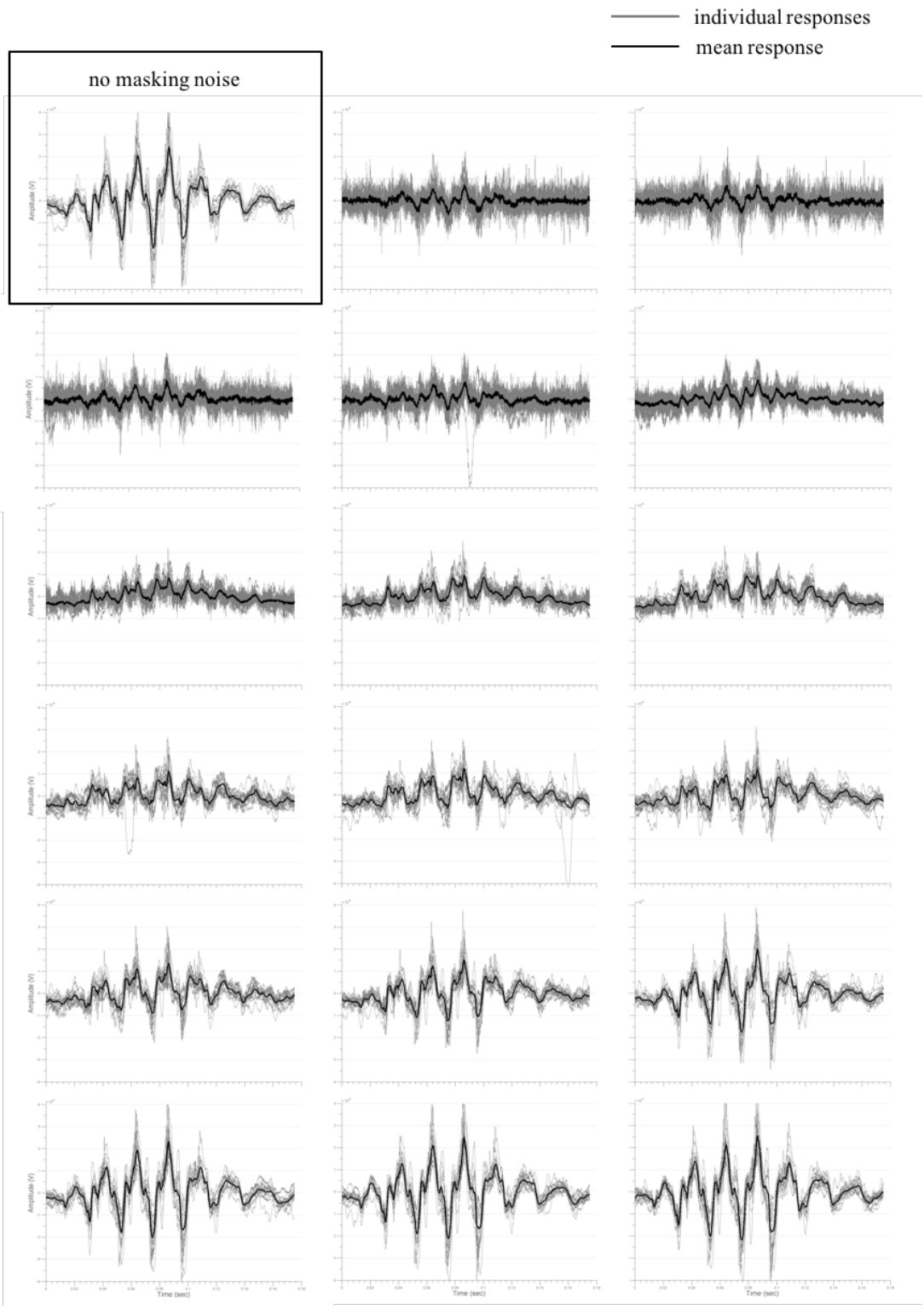

Figure 3. The 45 Hz, 80 dB SPL cochlear response waveform paneled by masker cutoff frequency, starting from the lowest cutoff frequency (146 Hz; top row, middle panel) to highest cutoff frequency (60 kHz; bottom row, right panel. Mean amplitude as a function of time is plotted as the solid black line, individual animals' waveforms are shown in gray.

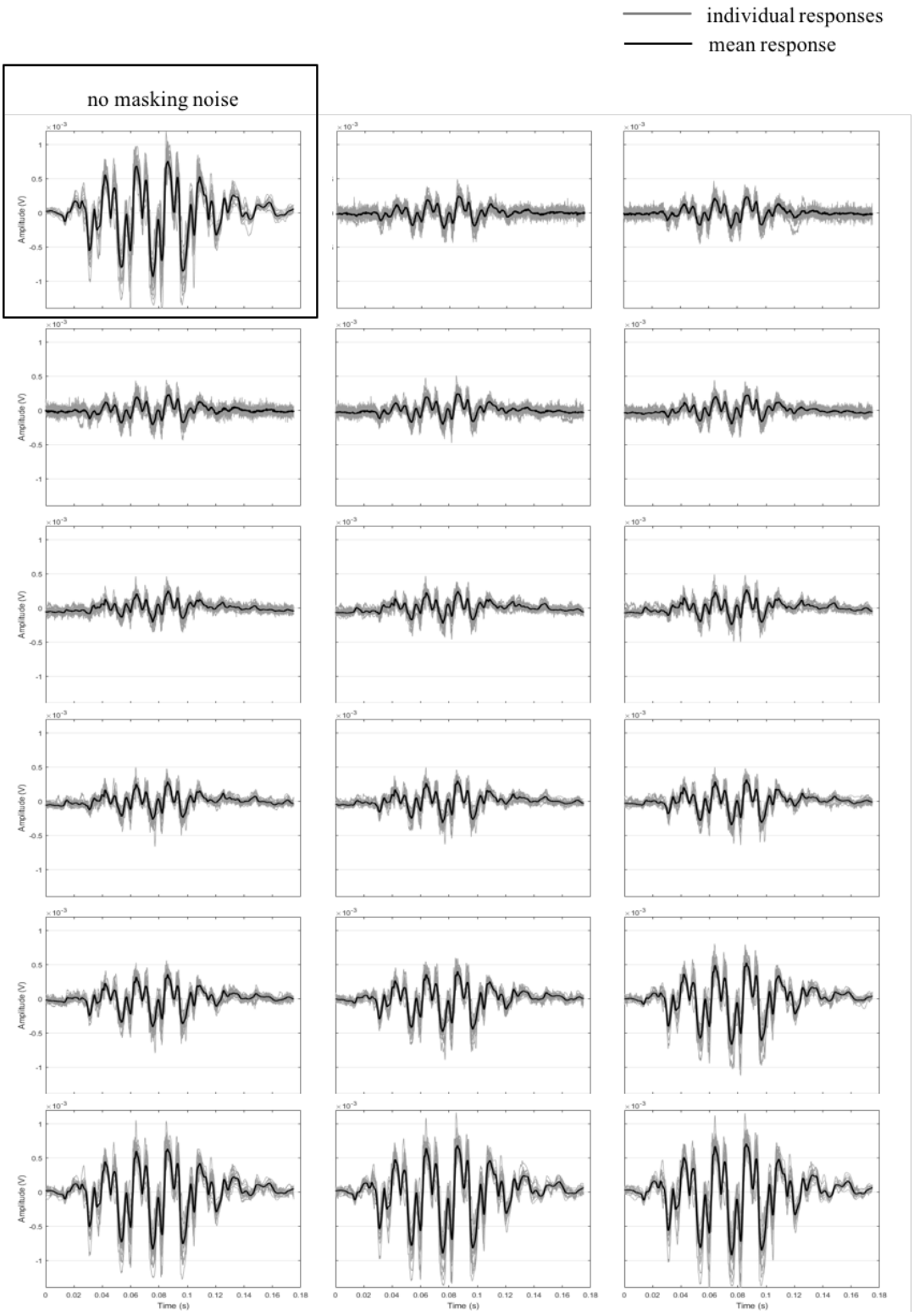

Figure 4. The 45 Hz, 90 dB SPL cochlear response waveform paneled by masker cutoff frequency, starting from the lowest cutoff frequency (146 Hz; top row, middle panel) to highest cutoff frequency (60 kHz; bottom row, right panel. Mean amplitude as a function of time is plotted as the solid black line, individual animals' waveforms are shown in gray.

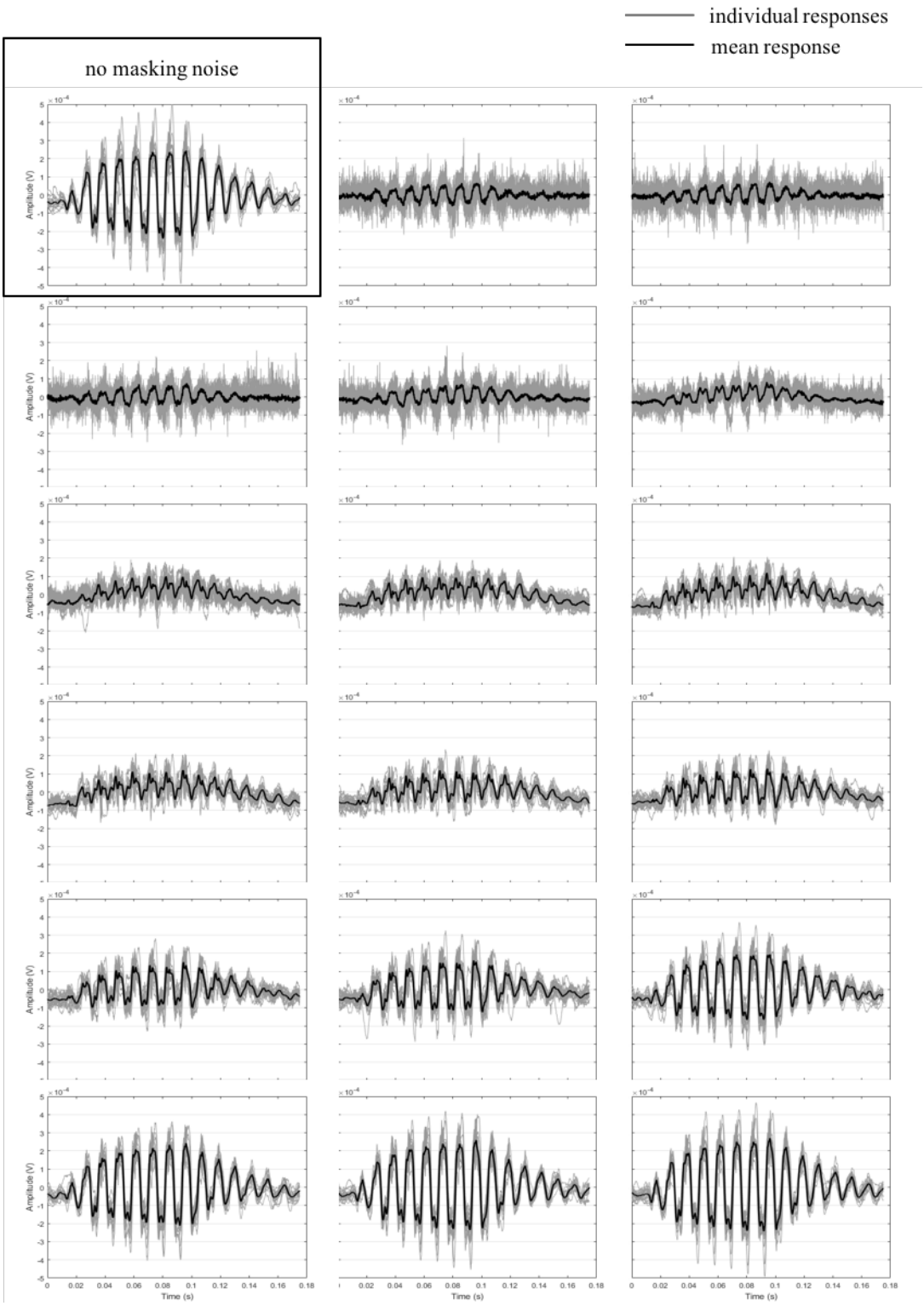

Figure 5. The 85 Hz, 80 dB SPL cochlear response waveform paneled by masker cutoff frequency, starting from the lowest cutoff frequency (146 Hz; top row, middle panel) to highest cutoff frequency (60 kHz; bottom row, right panel. Mean amplitude as a function of time is plotted as the solid black line, individual animals' waveforms are shown in gray.

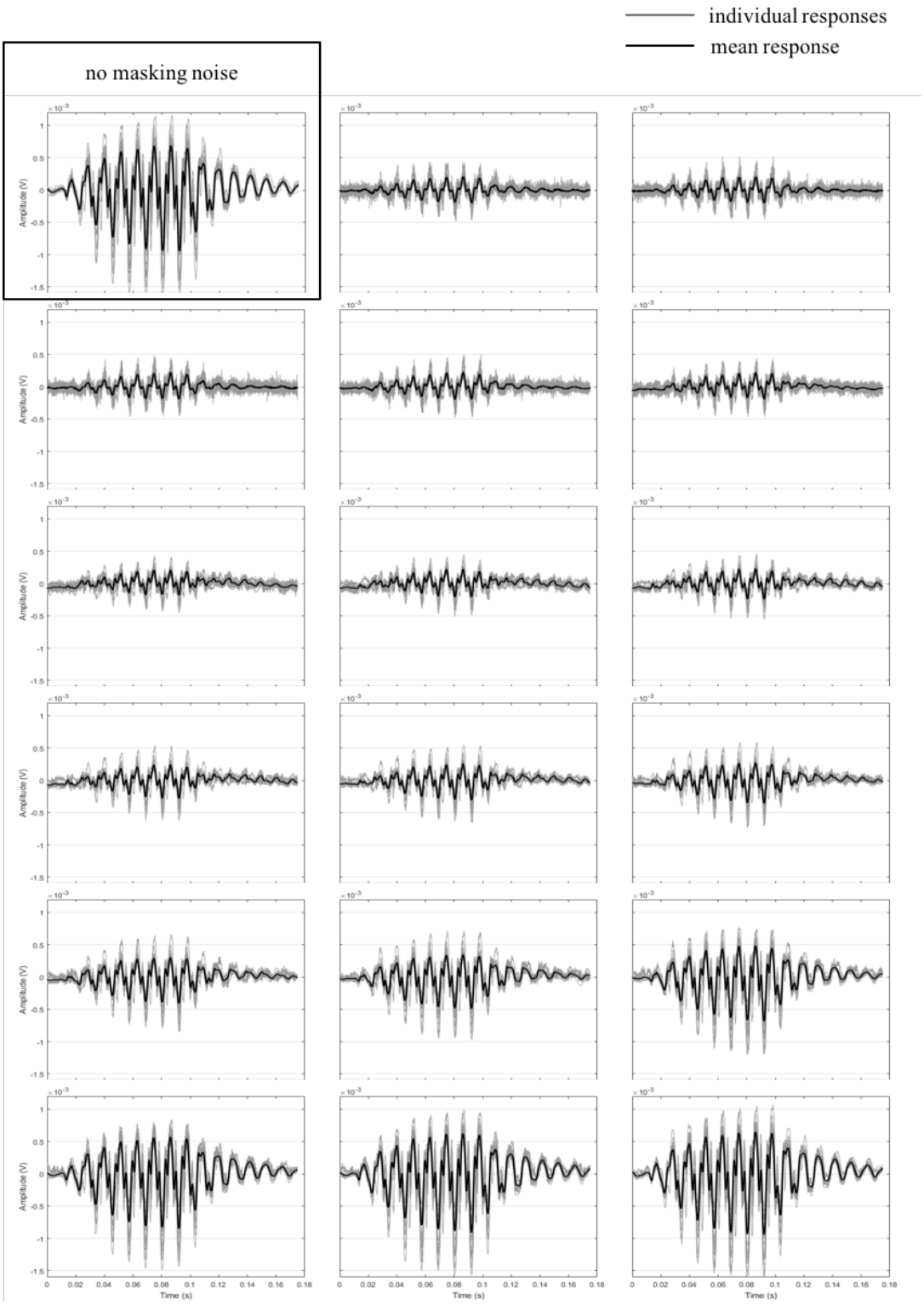

Figure 6. The 85 Hz, 90 dB SPL cochlear response waveform paneled by masker cutoff frequency, starting from the lowest cutoff frequency (146 Hz; top row, middle panel) to highest cutoff frequency (60 kHz; bottom row, right panel. Mean amplitude as a function of time is plotted as the solid black line, individual animals' waveforms are shown in gray.

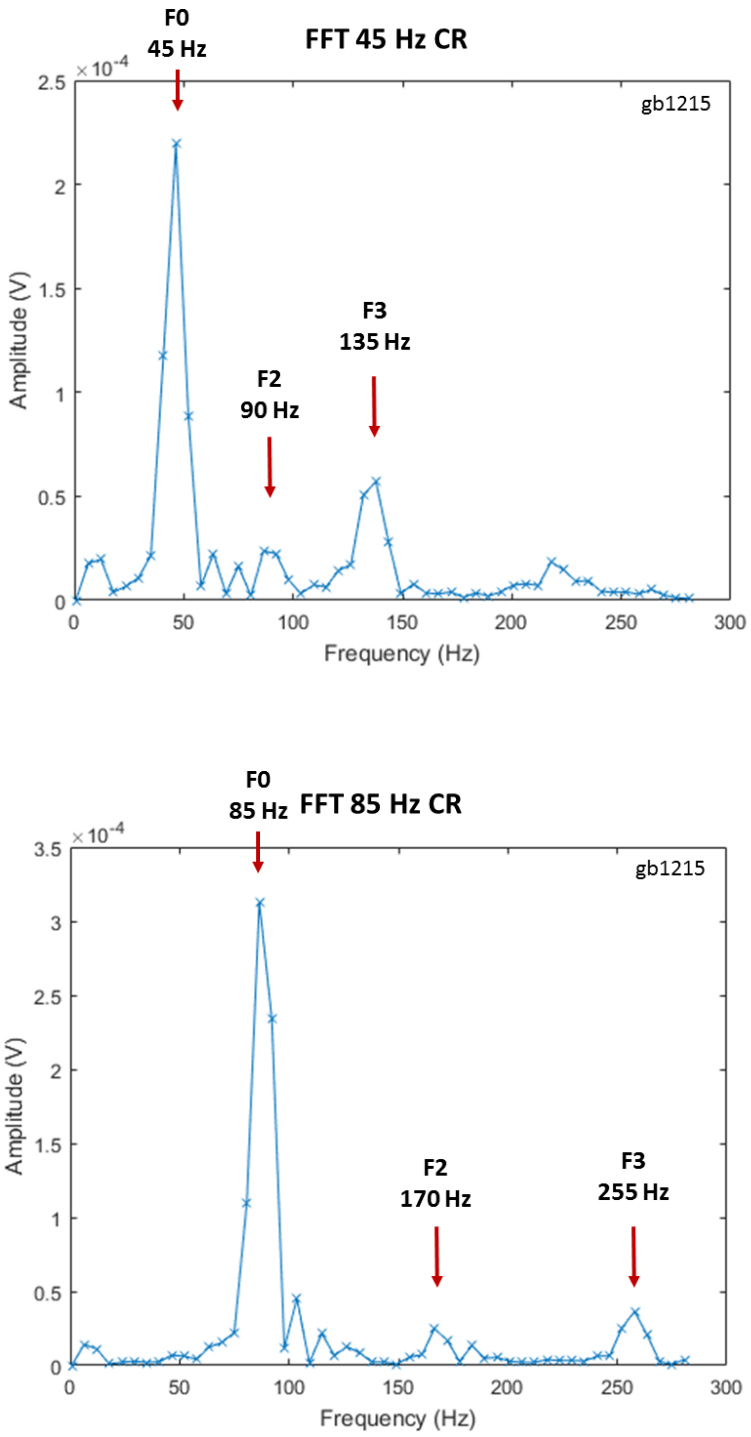

Figure 7. The FFT of one animal's 45 Hz (top) and 85 Hz (bottom) cochlear response. The amplitude of each harmonic (F2 & F3) found in the low-frequency CR was considered the magnitude of the peak (arrows).

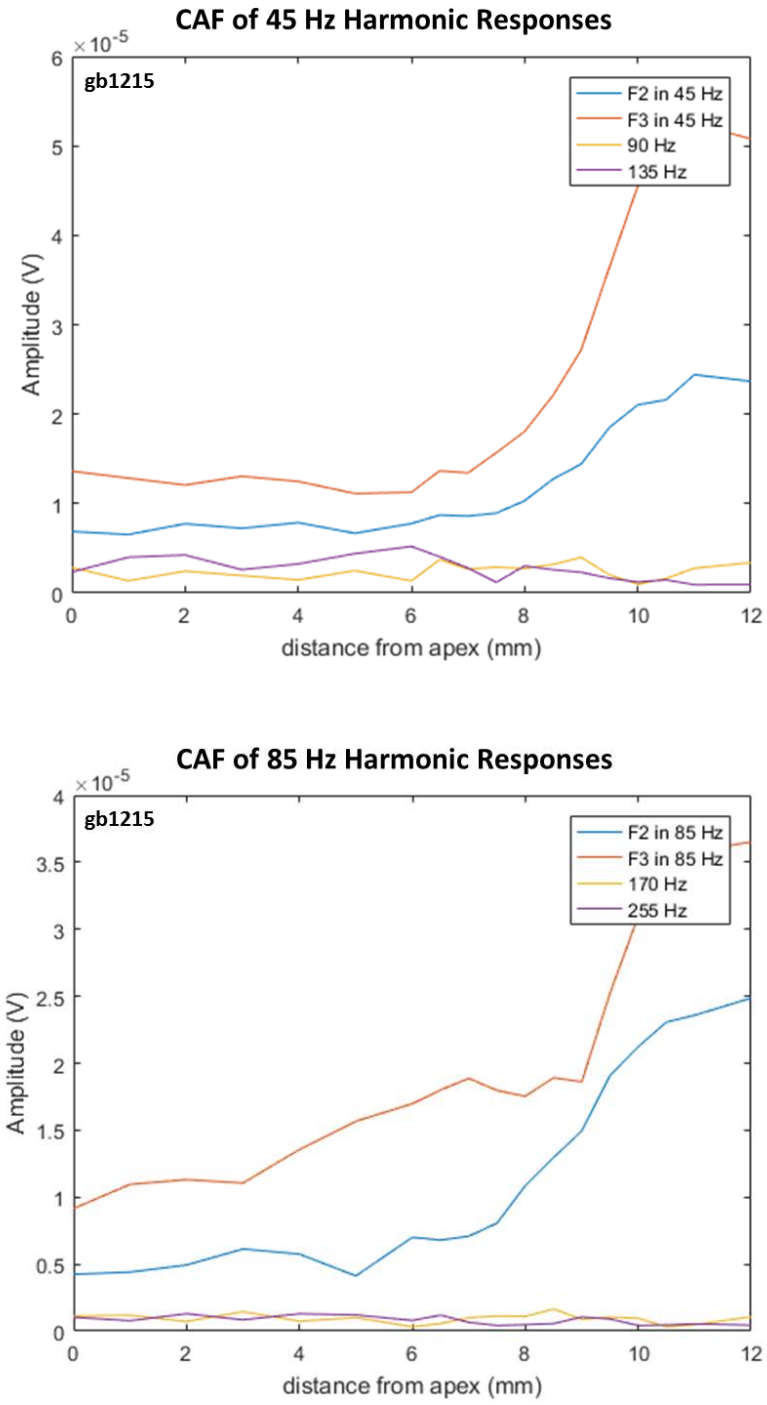

Figure 8. Demonstration of the fact that the CR distortion components are larger than a puretone response to the same frequency. The amplitude of the harmonics found in the animal's FFT of the cochlear response to 45 Hz (top) and 85 Hz (bottom) tones were plotted as a function of distance from the apex converted from masker cutoff frequency (orange and blue lines) were compared to the amplitude of the response to the individual harmonic stimuli generated to mirror the level of these harmonic distortions found in the acoustic stimulus (purple and yellow).

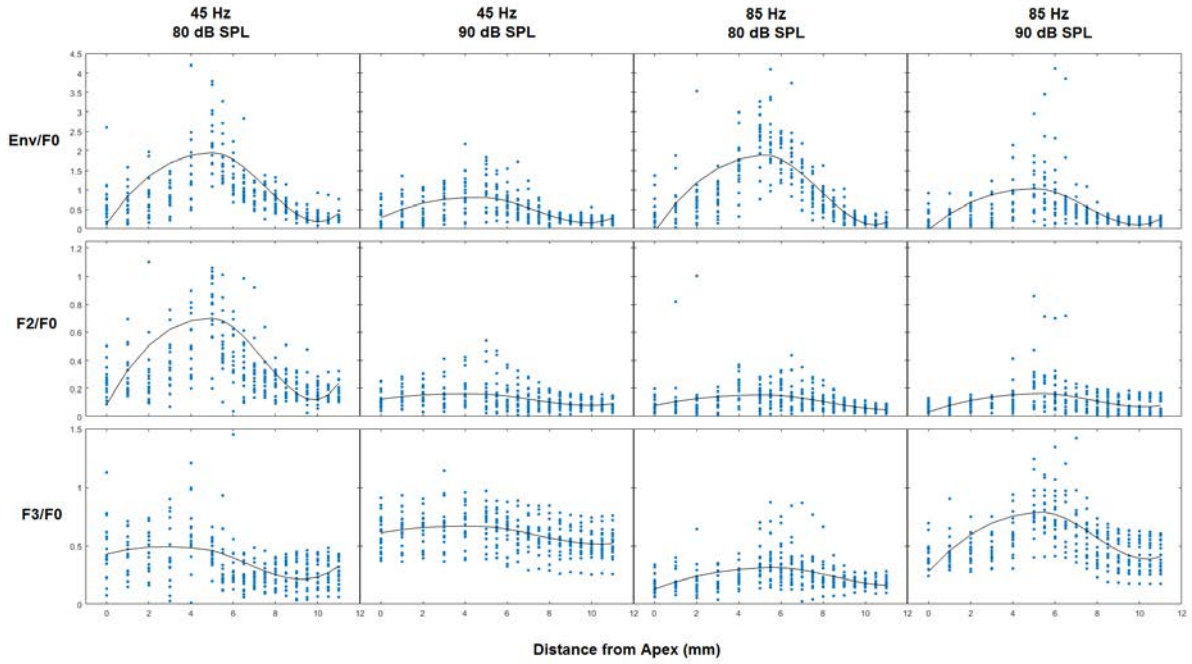

Figure 9. plots each harmonic ratio as a function of distance from the apex at each stimulus condition for all animals (blue dots). The data were regressed with harmonic as a covariate (solid black line) for each stimulus condition so that comparisons could be made across harmonic ratio. The analysis shows that the growth function of the Env/F0 (top row) is significantly different than F2 & F3 ratios at all stimulus conditions. The F2/F0 (middle row) and F3/F0 (bottom row) were significantly different at the 85 Hz, 90 dB SPL but not at other stimulus conditions.

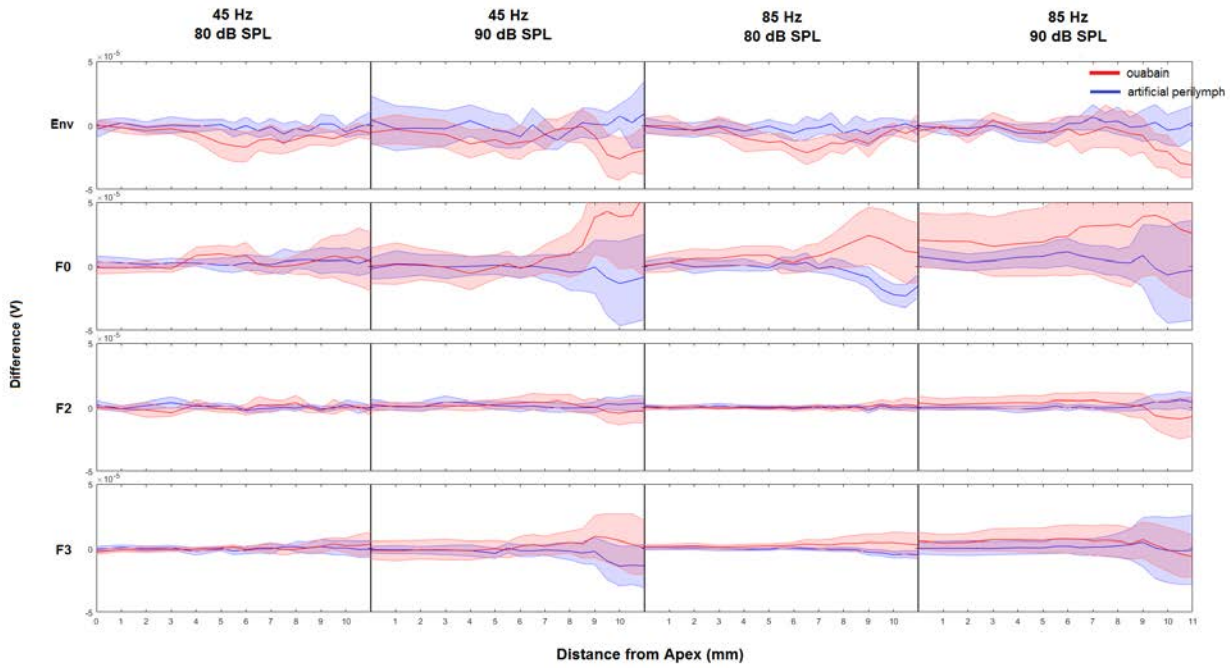

Figure 10. plots the mean (solid line) and 95% confidence intervals (shaded regions) for the difference between before and after treatment with either ouabain (red) or artificial perilymph (blue). Each row represents the difference data for one harmonic component and each column is a stimulus condition. The largest changes occur in the F0 component, where a slight decrease in the amplitude of the response at the base for animals treated with artificial perilymph whereas animals treated with ouabain show a slight increase in the F0 of the response from this region of the cochlear partition. The Env (top row) and F2 (middle row) components also showed a significant treatment group interaction with distance. The F3 (bottom row) component did not show any significant difference between animals treated with ouabain or artificial perilymph. The effect of ouabain was also dependent on stimulus condition, with no significant treatment effect at 45 Hz, 80 dB SPL (left column) but all other stimulus conditions showing treatment effects varying as a function of distance and harmonic component.

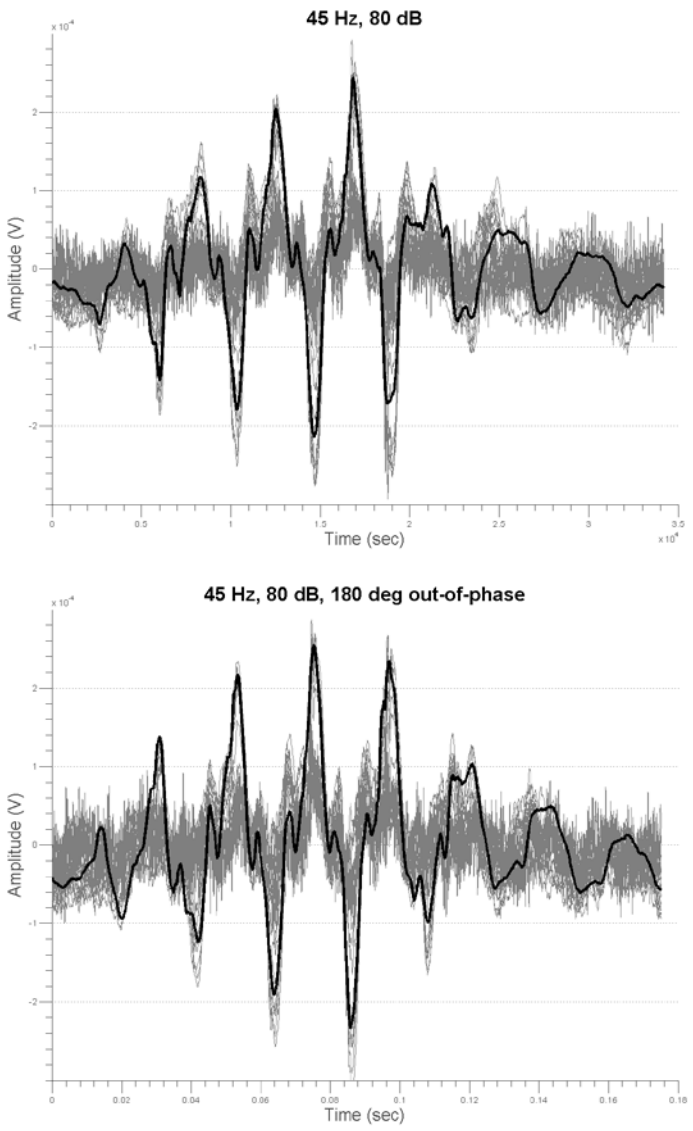

Figure 11. plots the cochlear response waveforms for the 45 Hz, 80 dB SPL rarefaction (top) and condensation stimuli (bottom). The mean (black) is plotted over the individual animals' waveforms (gray). These means were used as Mixtures 1 & 2 for the independent component analysis.

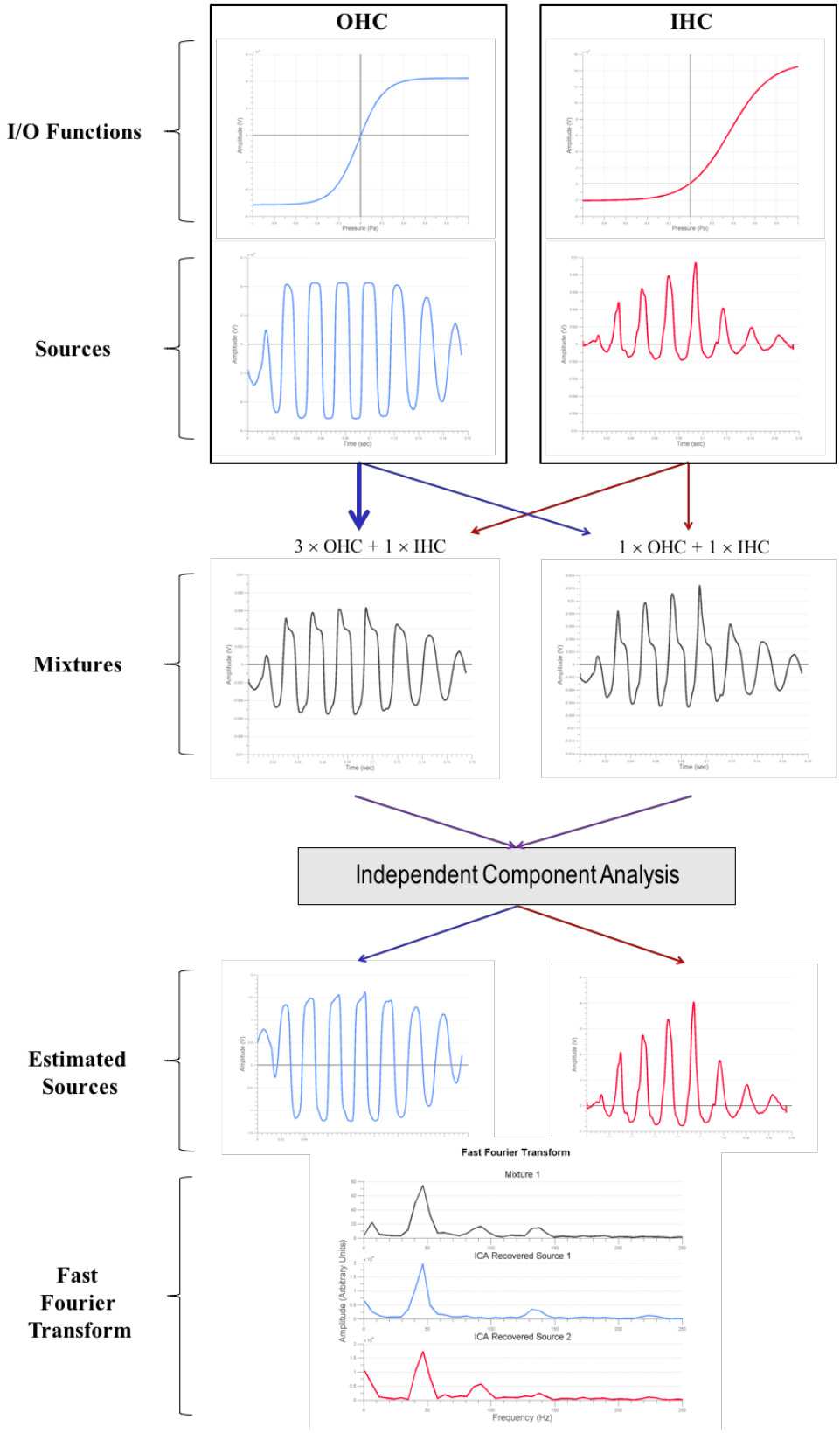

Figure 12. ICA on simulated data.

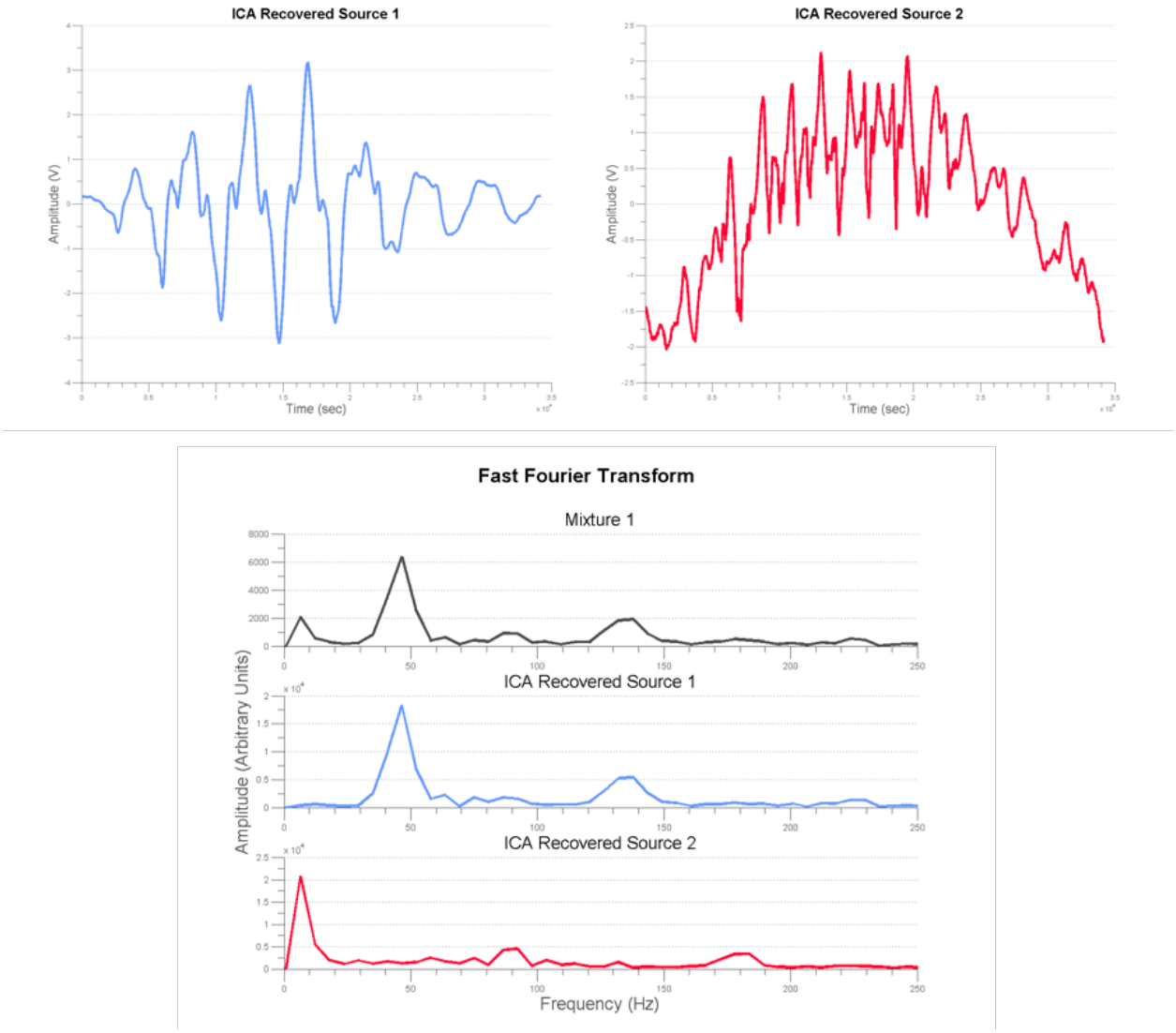

Figure 13. The top row plots the time-domain waveforms of the recovered sources derived from the real mean cochlear response to the 45 Hz, 80 dB SPL rarefaction and condensation stimuli. The amplitudes of the recovered sources are arbitrary due to the scaling process of ICA, but the spectral components contained by each recovered source (bottom; red and blue) can be used as evidence for identifying the cellular components.

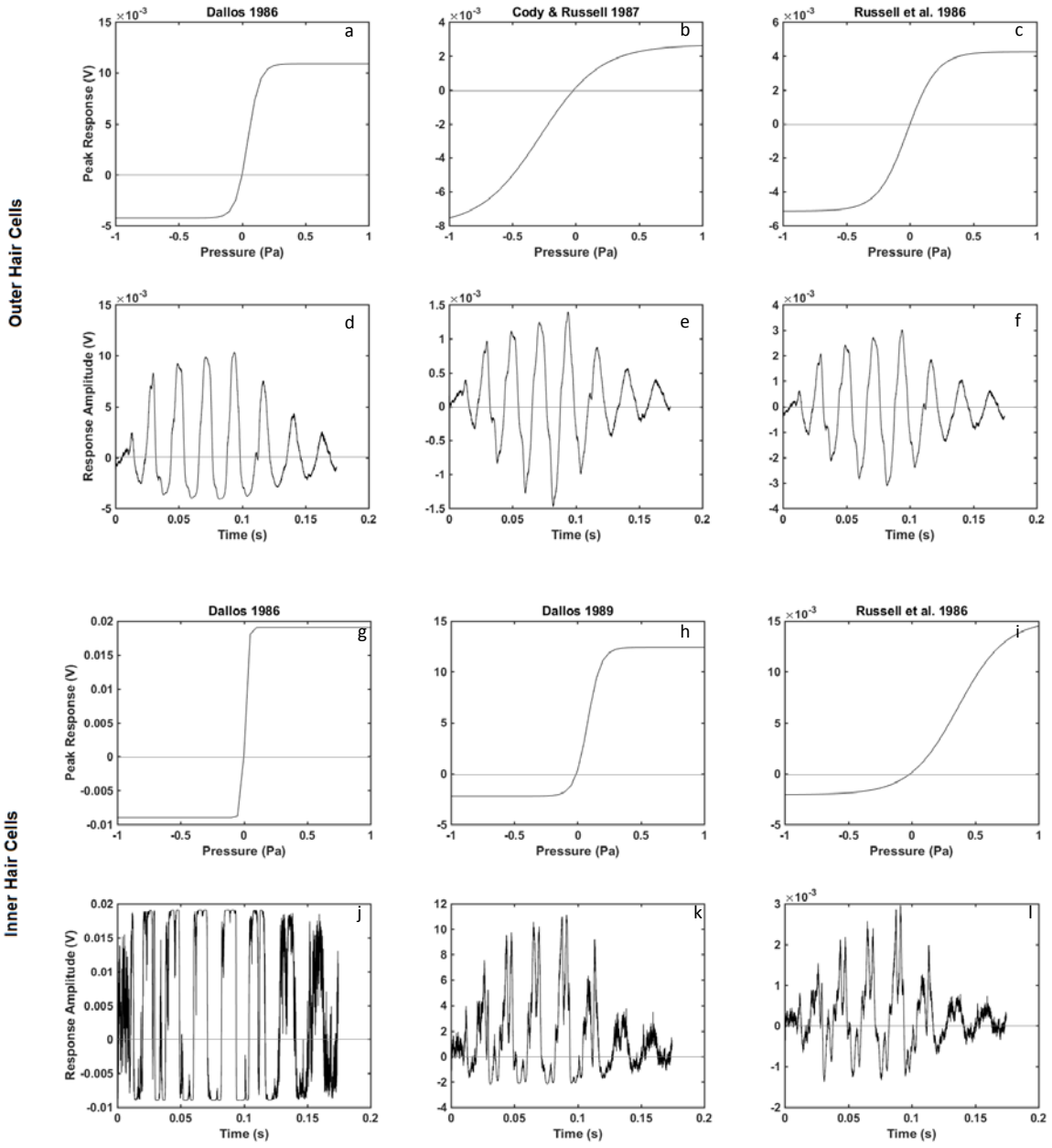

Figure 14. Transducer functions for outer hair cells (a-c) recreated from data by Dallos (1986), Cody & Russell (1987), and Russell, Cody, and Richardson (1986) and inner hair cells from Dallos (1986), Dallos (1989), and Russell, Cody, and Richardson (1986). The first order Boltzmann equation that fit each transducer curve was solved using appropriate derivation of the 45 Hz, 80 dB SPL stimulus for inner and outer hair cells and the resulting simulated response waveforms for each hair cell are plotted below their respective transducer curve in the second and fourth rows (d-f & j-l). The horizontal gray line shows where the response amplitude = 0, which clarifies any asymmetry in the transducer curves and the resulting DC shift in the response waveform.

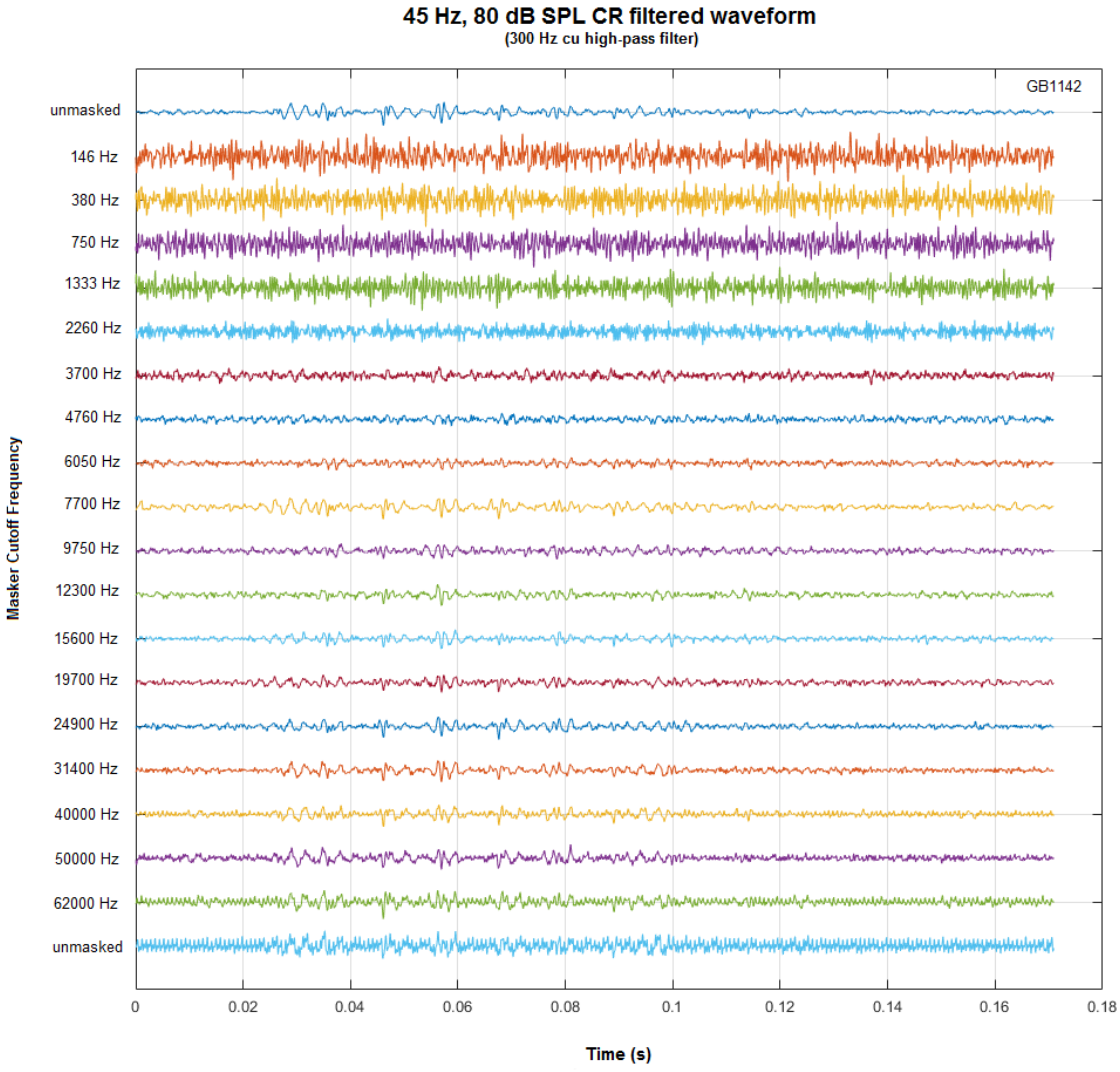

Figure 15. The phase-locked action potentials in the 45 Hz, 80 dB cochlear response waveform for one animal. Post-hoc high-pass filtering of the CR at 300 Hz attnuates the F0 and several orders of harmonics which highlights the phase-locked action potentials of the apical auditory nerve. The first row plots the APs in the unmasked condition and each consecutive row are the APs in the masking conditions from the lowest cutoff frequency (second row) to the highest (second-to-last row). The largest APs are located at y = ± .707 x peak for each cycle of the waveform (AP+ & AP-), but there are a number of APs phase-locked at various locations along the CR waveform.

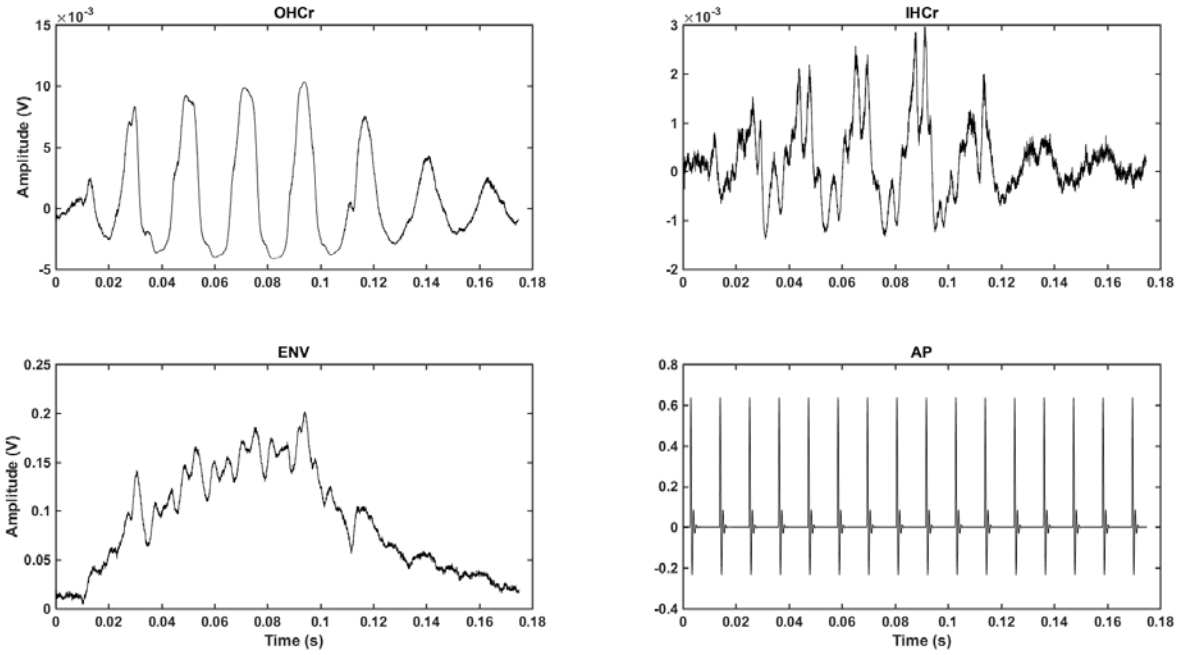

Figure 16. The covariates of the regression model in the time-domain.

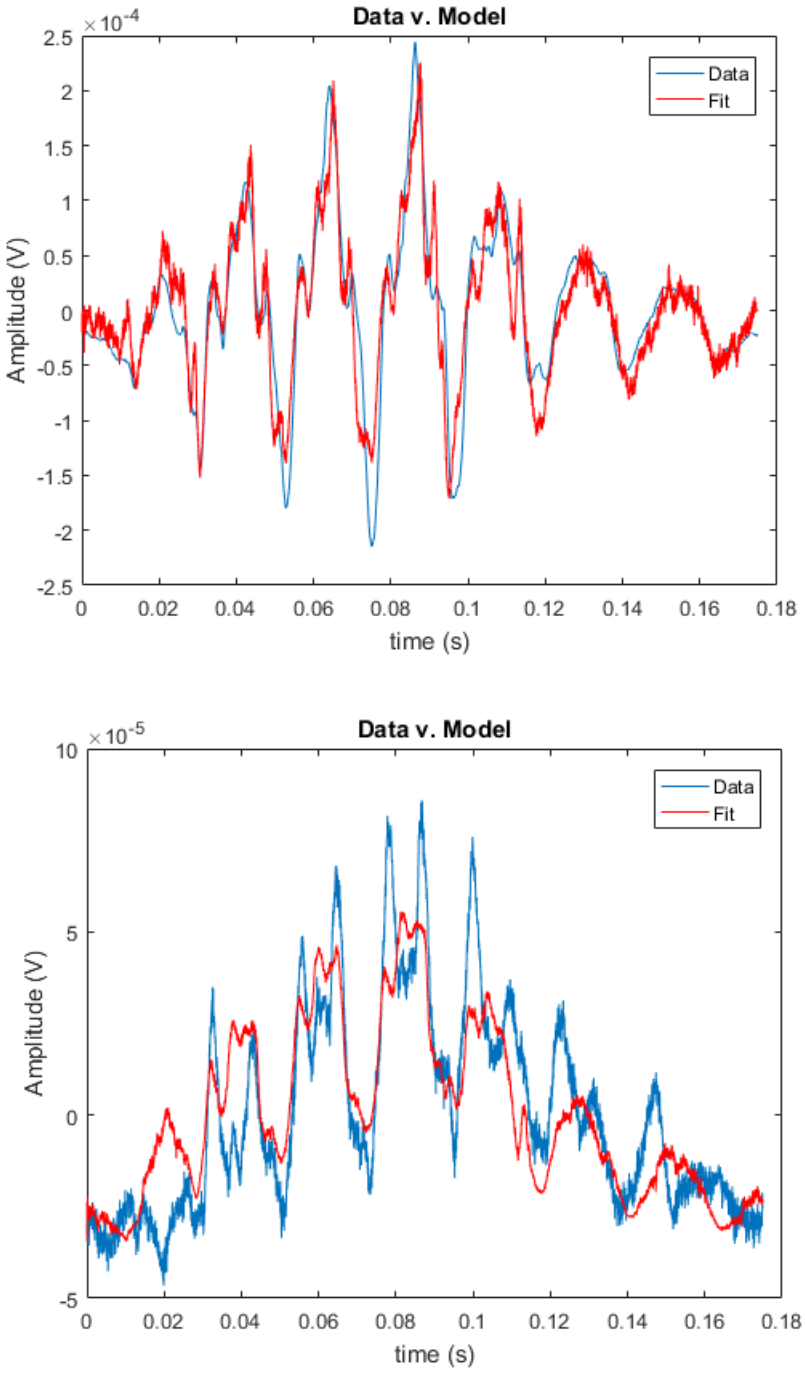

Figure 17. The estimated 45 Hz, 80 dB SPL cochlear response waveform (red) compared to the real data (blue) in the unmasked condition (top) and masker cutoff of 3700 Hz (bottom).

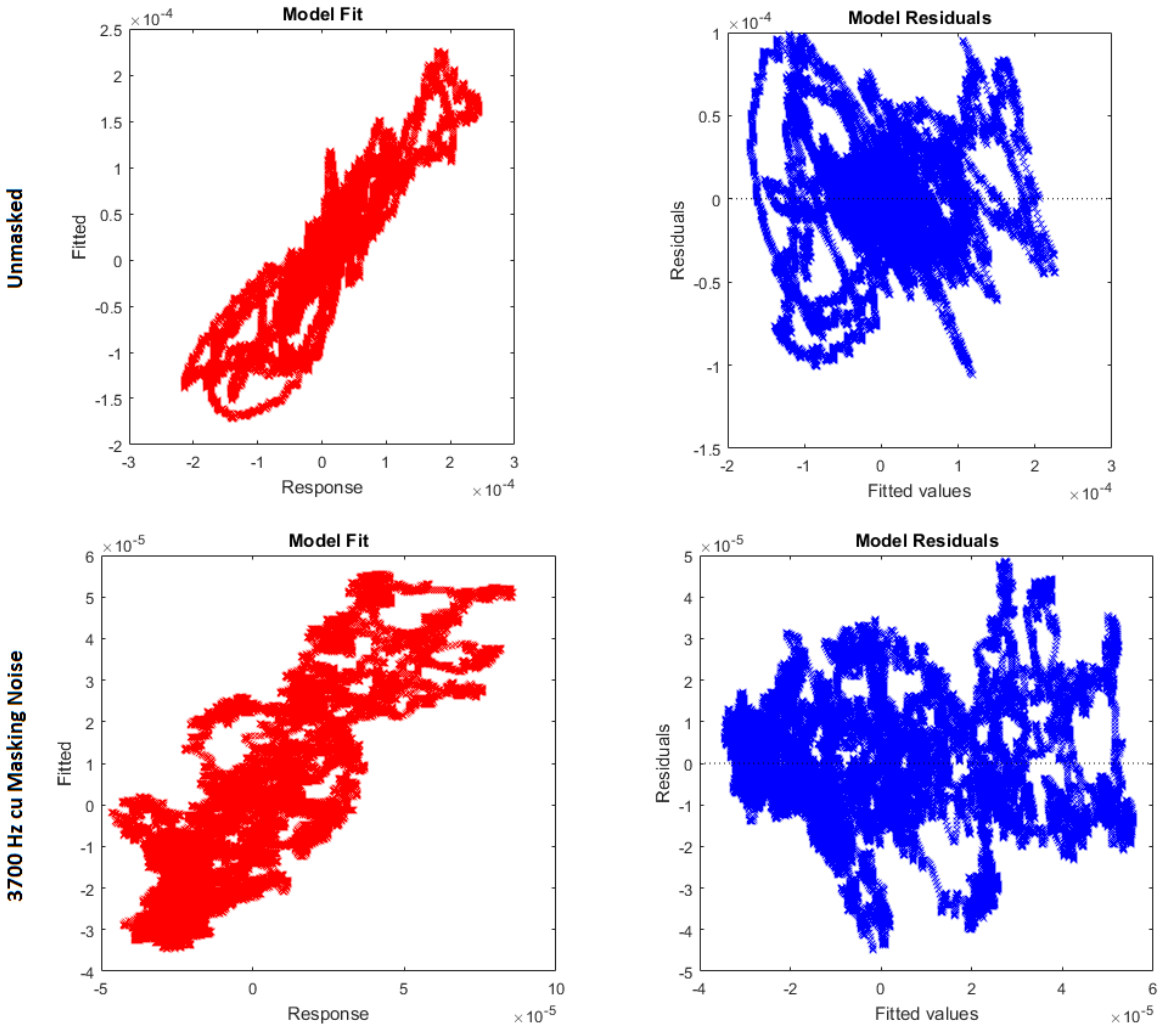

Figure 18. The model fit (left panels) for the unmasked 45 Hz, 80 dB SPL CR (top) and masking condition 6 (3700 Hz cutoff frequency; bottom) which plot the estimated values against the real data values, showing a good correlation between the model CR and the data. The residuals (right panels) are centered around zero implying the model was an appropriate fit to the data.

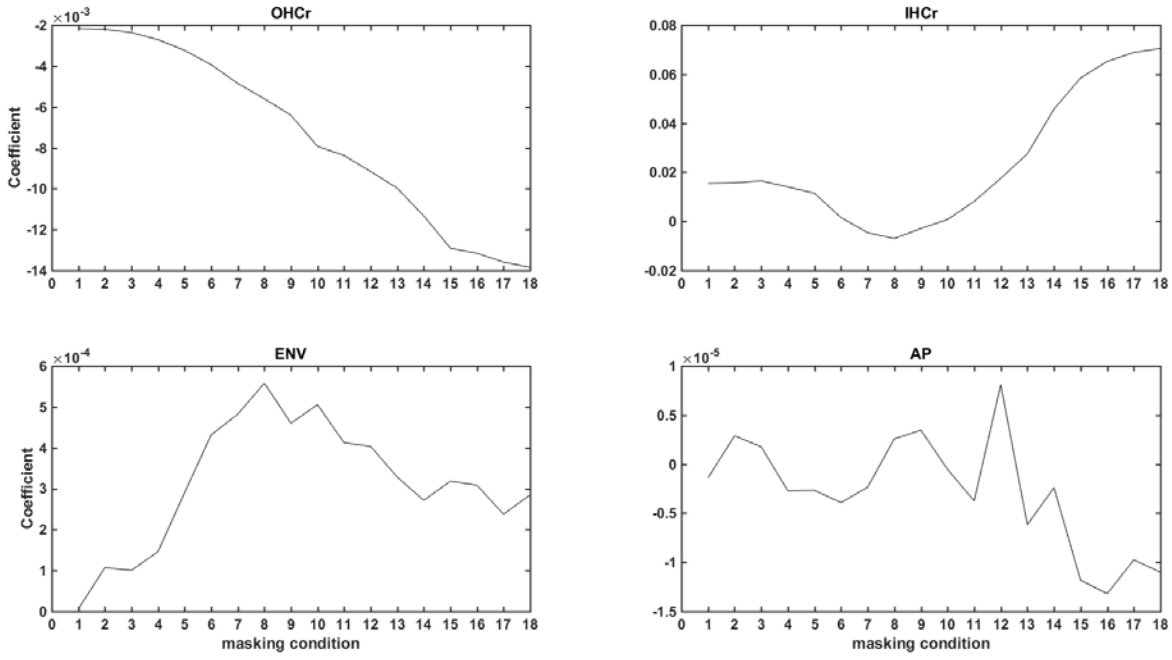

Figure 19. The estimated coefficients for each covariate as a function of masking condition from lowest cutoff frequency to highest.

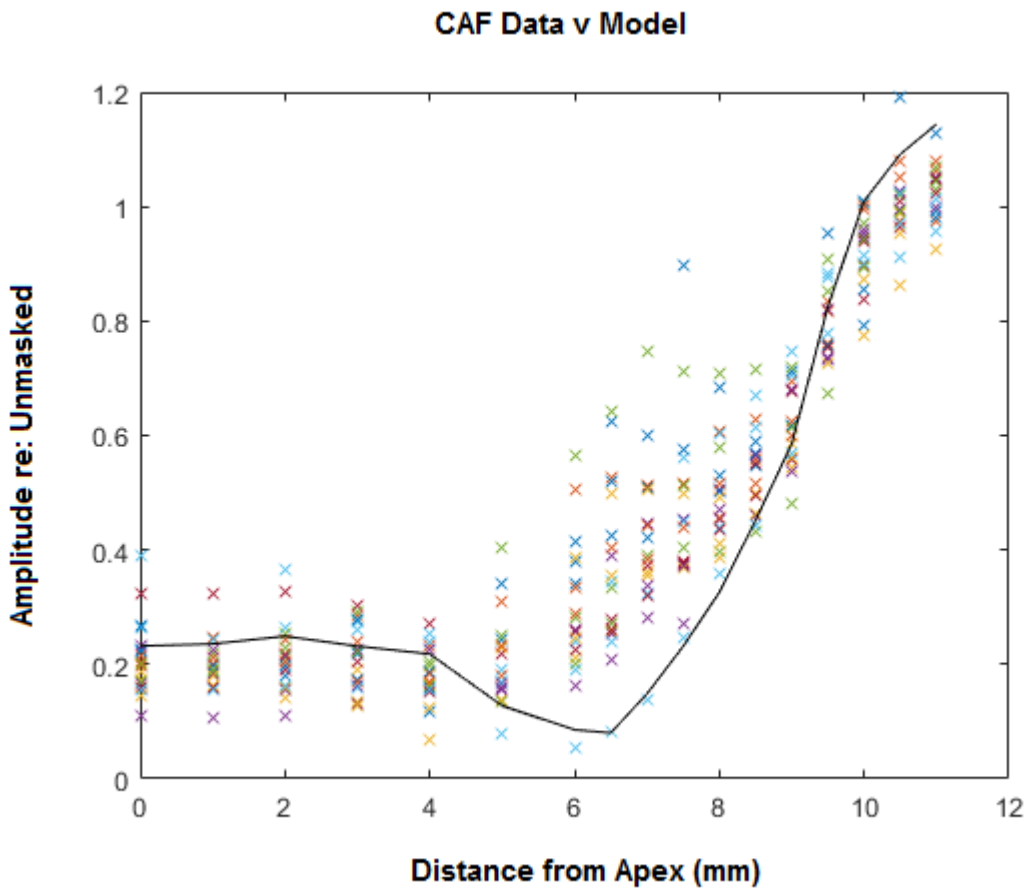

Figure 20. The estimated 45 Hz, 80 dB SPL CAF using the magnitude and phase of F0 from the *OHCr* and *IHCr* covariates (solid line) plotted alongside the real data from all animals (n=18).

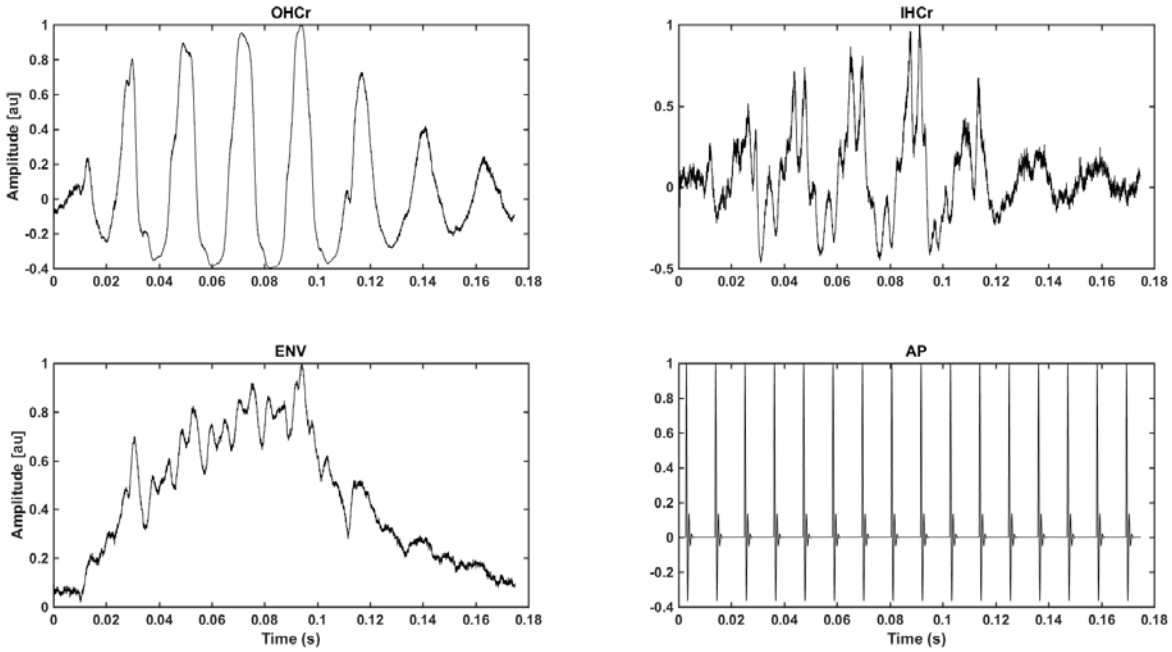

Figure 21. The covariates of the regression model in the time-domain normalized to their maximum amplitudes.

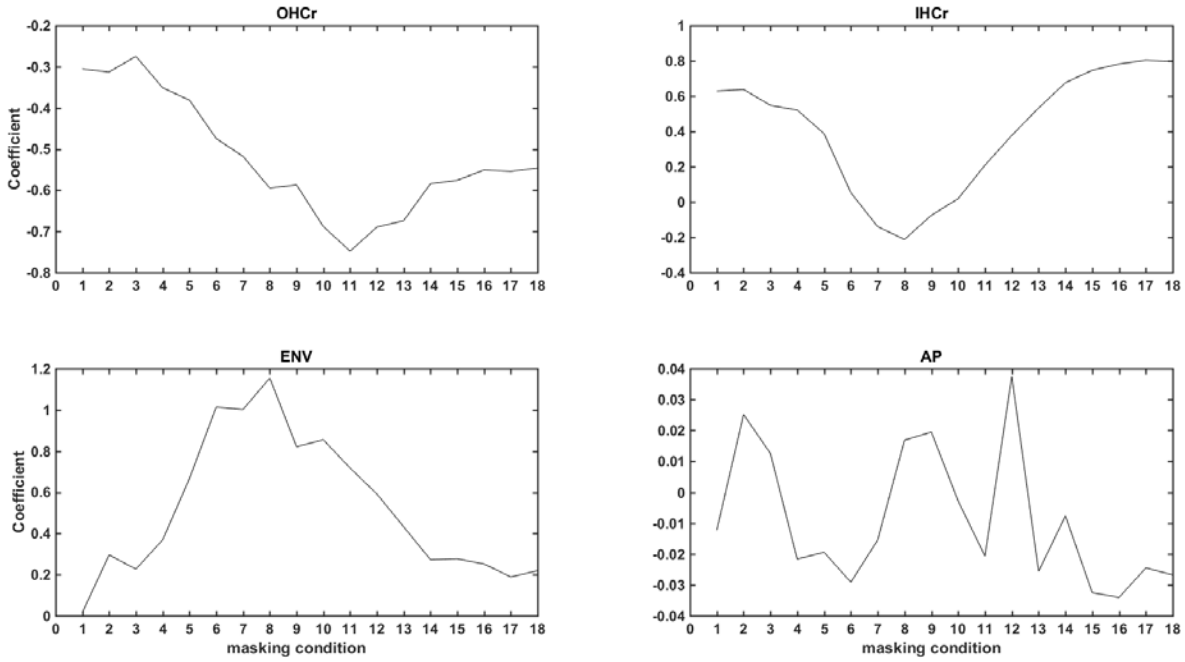

Figure 22. The estimated coefficients of the normalized covariates as a function of masking condition from lowest cutoff frequency to highest. At the low cutoff frequencies the *IHCr* contribution is up to three times larger than the *OHCr*; in the mid-cutoff frequencies, the *IHCr* contribution is smaller than the *OHCr*; and at the highest cutoff frequencies the two contributions are relatively equal.

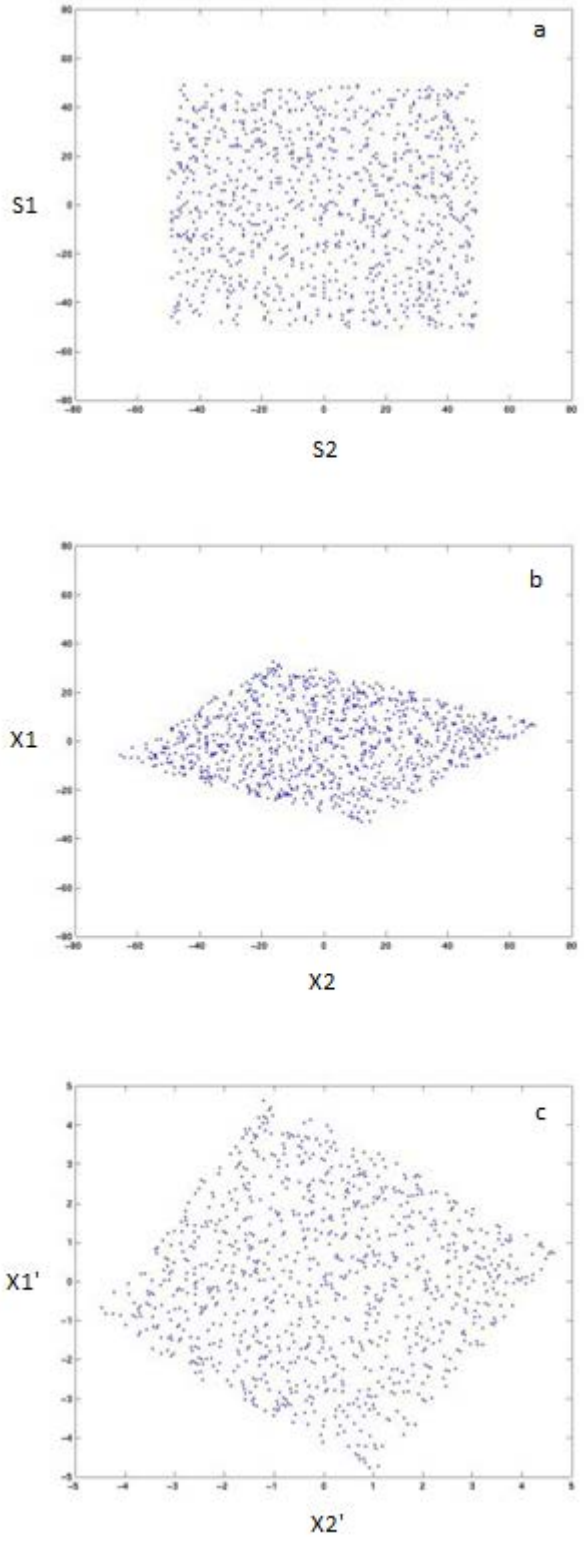

Figure 23. Plots of two example source amplitudes (a); mixtures amplitudes (b); and whitened mixture amplitudes(c).

Table 1. Estimated coefficients and 95% confidence intervals for the main effect of each harmonic ratio as a function of the distance variables of the third-order polynomial. The bottom row of each stimulus condition reports the significance of harmonic on the growth curve of the ratios re: F0. A significant p-value for any distance variable indicates an effect of harmonic of the shape of the growth curve, while a significant main effect indicates a constant amplitude difference between at least two harmonic ratios. * indicates a significance of $p < .05$, and ** a significance of $p < .01$.

		α **main effect**	β_1 **linear distance**	β_2 **quadratic distance**	β_3 **cubic distance**
45 Hz 80 dB SPL	*Env/F0*	0.137 (-0.292,0.565)	0.806 (0.588,1.024)	-0.106 (-0.137,-0.076)	0.004 (0.002,0.005)
	F2/F0	0.089 (-0.340,0.518)	0.278 (0.060,0.497)	-0.037 (-0.068,-0.007)	0.0012 (0.0001,0.0024)
	F3/F0	0.428 (-0.0005,0.857)	0.050 (-0.168,0.268)	-0.011 (-0.414,0.019)	0.0005 (-0.0007,0.001)
	p-value	*0.417*	*0.000***	*0.000***	*0.000***
45 Hz 90 dB SPL	*Env/F0*	0.288 (0.212,0.365)	0.252 (0.219,0.285)	-0.036 (-0.040., 0.032)	0.0013 (0.0011,0.0014)
	F2/F0	0.125 (0.0485,0.202)	0.022 (-0.011,0.055)	-0.004 (-0.008,0.001)	0.0001 (-0.00004,0.0003)
	F3/F0	0.609 (0.532,0.685)	0.032 (-0.001,0.065)	-0.005 (-0.01,-0.0006)	0.0002 (0,0.0004)
	p-value	*0.001***	*0.000***	*0.000***	*0.000***
85 Hz 80 dB SPL	*Env/F0*	-0.064 (-0.193,0.065)	0.814 (0.748,0.880)	-0.100 (-0.110,-0.091)	0.003 (0.002,0.004)
	F2/F0	0.278 (-0.047,0.602)	0.164 (0.010,0.319)	-0.024 (-0.046,-0.003)	0.0009 (0.00003,0.002)
	F3/F0	0.132 (0.003,0.261)	0.067 (0.0004,0.132)	-0.007 (-0.016,0.002)	0.0002 (-0.0002,0.0006)
	p-value	*0.054*	*0.000***	*0.000***	*0.000***
85 Hz 90 dB SPL	*Env/F0*	-0.022 (-0.180,0.136)	0.463 (0.391,0.536)	-0.608 (-0.701,-0.051)	0.002 (0.0016,0.0024)
	F2/F0	0.080 (-0.050,0.209)	0.032 (-0.034,0.098)	-0.004 (-0.013,0.005)	0.0001 (-0.0002,0.0004)
	F3/F0	0.276 (0.119,0.433)	0.206 (0.134,0.280)	-0.025 (-0.035,-0.014)	0.0008 (0.0004,0.0012)
	p-value	*0.020**	*0.000***	*0.000***	*0.000***

Table 2. The t-values (top) and p-values (bottom) of the pairwise comparisons between harmonic ratios as a function of the distance variables. A significant p-value indicates a difference in the growth curves between the harmonic ratios.
* indicates a significance of $p < .05$, and ** a significance of $p < .01$.

		d interaction with		**d^2** interaction with		**d^3** interaction with	
		F2/F0	*F3/F0*	*F2/F0*	*F3/F0*	*F2/F0*	*F3/F0*
45 Hz 80 dB SPL	*Env/F0*	-3.354 0.000**	-4.804 0.000**	3.140 0.002**	4.351 0.000**	-2.650 0.008**	-3.620 0.008**
	F2/F0		-1.450 0.147		1.211 0.226		-0.970 0.332
45 Hz 90 dB SPL	*Env/F0*	9.645 0.000**	-9.217 0.000**	9.847 0.000**	9.361 0.000**	-8.813 0.000**	-8.440 0.000**
	F2/F0		0.428 0.669		-0.486 0.627		0.373 0.709
85 Hz 80 dB SPL	*Env/F0*	-16.376 0.000**	-15.658 0.000**	14.539 0.000**	14.041 0.000**	-11.796 0.000**	-11.453 0.000**
	F2/F0		.718 .473		-.498 .619		.344 .731
85 Hz 90 dB SPL	*Env/F0*	-7.818 0.000**	-4.905 0.000**	7.451 0.000**	4.926 0.000**	-6.478 0.000**	-4.452 0.000**
	F2/F0		2.913 0.004**		-2.524 0.012*		2.026 0.043*

Table 3. The R^2 values of the regression model fit to the real 45 Hz, 90 dB SPL cochlear response using each combination of possible simulated inner (*IHCr*) and outer hair cells (*OHCr*). The bolded values indicate the best fit and two hair cells used in the final model.

		OHC					
		Dallos (1986)		Cody & Russell (1987)		Russell et al. (1986)	
		unmasked	*masked*	*unmasked*	*masked*	*unmasked*	*masked*
IHC	Dallos (1986)	0.68	0.70	0.65	0.63	0.67	0.65
	Dallos & Cheatham (1989)	0.76	0.70	0.76	0.63	0.77	0.65
	Russell et al. (1986)	**0.81**	**0.70**	0.79	0.63	0.80	0.65