KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Neuroscience
Laboratory of Auditory Neurophysiology

KU LEUVEN

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

Assessment of Fundamental Cochlear Limits of Frequency Resolution and Phase-locking in Humans and Animal Models

Eric VERSCHOOTEN

Promoter:
Prof. dr. Philip Joris

Dissertation presented
in partial fulfillment of the
requirements for the
degree of Doctor in
Biomedical Sciences

July 2013

Acknowledgements

This thesis is the result of many years hard work at the Laboratory of Auditory Neurophysiology.
It would not have been possible without the efforts and help of many people, to whom I am sincerely grateful.

Firstly, I would like to thank my promoter, Prof. Philip Joris who was abundantly helpful and offered invaluable assistance and support and guidance.

I wish to express my gratitude to the staff and colleagues of the department of Neuroscience for their administrative support and to allow auditory testing of several macaque monkeys. Also to doctor Prof. Christian Desloovere for placing the needle electrodes during my human experiments. And especially to Prof. Luis Robles, with whom I conducted chinchilla and cat experiments during his visit in the summer periods.

I am grateful to the members of the jury, Prof. Ian Winter, Prof. Stefaan Peeters, Prof. Sabine Van Huffel and Prof. Jan Wouters for taking the time to evaluate this thesis.
I also wish to thank all my (ex-) colleagues at the Laboratory of Auditory Neurophysiology who surrounded me with their joy and enthusiasm: Damir Kovacic, Myles Mc Laughlin, Peter Bremen, Pascal Michelet, Kevin Spiritus, Shotaro Karino, Abel Jonckheer, Tom Franken, Tina Wei, …

A special word of thanks I want to address to Prof. Stefaan Peeters, who believed in my skills and recommended me to Prof. Philp Joris for a PhD in auditory neuroscience.

Finally, I express my deepest gratitude to my family and especially to my wife Bianca and daughter Zoë, for their *patience!*, understanding and support during these years of devotion to science and research.

Jury

Promoter

Prof. Dr. Philip Joris	University of Leuven

Chair

Prof. Dr. Mark Jorissen	University of Leuven

Doctoral Jury members

Prof. Dr. Ian Winter	University of Cambridge
Prof. Dr. ir Stefaan Peeters	University of Antwerp
Prof. Dr. ir Sabine Van Huffel	University of Leuven
Prof. Dr. Jan Wouters	University of Leuven

.

Table of Contents

List of abbreviations xi

Abstract xiii

Chapter 1. General Introduction 1

Chapter 2. Auditory nerve frequency tuning measured with forward-masked compound action potentials 23

Chapter 3. Electrophysiological evidence for sharp cochlear tuning in humans 59

Chapter 4. Upper frequency limit of phase-locking in stimulus evoked potentials 73

Chapter 5. Upper frequency limit of phase-locking in human and monkey 121

Chapter 6. General Discussion 143

References 153

Summary/samenvatting 159

Bibliography 167

List of abbreviations

µs	microsecond
AC	alternating current
AN	auditory nerve
ANF	auditory nerve fiber
ANOW	auditory nerve overlapped waveform
BM	basilar membrane
BW	bandwidth
CAP	compound action potential
CF	characteristic frequency
CM	cochlear microphonic
CNS	central nervous system
C_v	coefficient of variation
dB	decibel
DC	direct current
ERB	equivalent rectangular bandwidth
FFT	fast Fourier transformation
FIR	finite impulse response
FTC	frequency-threshold tuning curve
GMB	growth-of-maskability
GOM	growth-of-masking
I/O	input-output
IHC	inner hair cell
L_x	level x
ms	millisecond
MTC	masking tuning curve
NNFM	notched-noise forward masking
NP	auditory nerve neurophonic
NW	notch-width
OHC	outer hair cell
PSTH	post-stimulus time histograms
Q	Quality factor
ROEX	rounded exponential
RW	round window
SEE	standard error of the estimate for regression
SEM	standard error of the mean
SFOAEs	stimulus-frequency otoacoustic emissions
SNfR	signal-to-noise floor ratio
SNR	signal-to-noise ratio
SPL	sound pressure level
STD	standard deviation
STFT	short-time Fourier transform
TTX	tetrodotoxin
VS	vector strength

ABSTRACT

Background

Frequency selectivity and neural phase-locking are two primary properties of the peripheral auditory system and are fundamental to hearing. These properties are reflected in neural signals as a place-rate and time code and are quantified by means of responses from single auditory nerve fibers (AN fibers). Because of practical issues, the limits of both properties are poorly characterized in humans and widely differing values are assumed in the interpretation of non-electrophysiological results. Recent otoacoustic and behavioral measurements suggest that frequency selectivity is higher in humans than in laboratory animals (Shera et al., 2002), but this is disputed based on comparisons of behavioral and electrophysiological measurements across species (Ruggero and Temchin, 2005). Neural phase-locking declines with frequency and becomes undetectable at an upper frequency limit which differs between species. Some studies suggest that phase-locking in human extends to much higher frequencies than that found in common laboratory animals such as cat. In this thesis we investigate the use of mass potentials to electrophysiologically quantify these properties in humans.

Methods

A minimally invasive transtympanic protocol was developed to record stimulus-evoked mass-potentials from the cochlear promontory or from the niche of the round window in monkey (macaque) and human. This involved a custom made ear mold with openings for the transtympanic needle electrode and for acoustic stimulation (ER-2 earphone, *in situ* calibrated with an ER-7 microphone). To obtain frequency selectivity a notched-noise forward masking paradigm (Oxenham and Shera, 2003) was combined with the recording of compound action potentials (CAP). For neural phase-locking, we developed a method based on forward masking to disambiguate contributions of receptor and neural origin. Both methods were first developed, assessed and validated at the round window in cat and thereafter applied in monkey and human.

Results

Frequency selectivity. We found that the empirical method that was first developed and assessed in cat was a suitable means to examine frequency tuning for a number of reasons. a) Masking of CAP responses behaved as expected from the physiology of single AN-fibers. b) Masking tuning curves were equivalent to tuning curve of single AN-fibers. c) The sharpness of masking tuning curves, quantified as a Q_{10}-factor, followed closely the lower boundary of single AN-fibers and was only slightly higher at high frequencies. CAP-tuning was dependent on probe level consistent with cochlear physiology. The CAP responses behaved similar in human and monkey, but the responses and signal-to-noise(floor) ratios were much smaller and therefore higher stimulus levels (>30 dB) were needed in human and monkey. We found that CAP tuning in human was on average 1.6 times sharper than in cat and chinchilla and 1.35 times sharper than in monkey (macaque).

Neural phase-locking. We found that potentials near the round window have a considerable neural phase-locked contribution that can be isolated and quantified with the present method. Moreover, the neural component was in many ways similar to that measured on the auditory nerve. The only noteworthy differences with the latter were a typically larger magnitude and a smaller time lag for the neural component at the round window. The frequency limit obtained in cat was 4.7 kHz, which is very similar to the limit reported for individual AN-fibers (~ 5 kHz). We found in many respect similarities between the measurements in human, monkey and those at the round window in cat, with similar temporal and spectral properties. The most notable difference in human and monkey compared to cat was a much smaller magnitude (-30 dB) and also signal-to-noise(floor) ratio for the neurophonic. The upper detectable frequency limits of neural phase-locking obtained from stimulus evoked potentials was 3.3 kHz in human and 4.0 kHz in monkey (macaque).

Conclusions

Frequency selectivity and neural phase-locking can be characterized from mass potentials with a minimally invasive technique from awake, normal hearing subjects.
We found evidence that a) humans have sharper frequency tuning than commonly studied animals, consistent with previous reports (Shera et al., 2002; Oxenham and Shera, 2003) and; b) humans have an upper frequency limit of phase-locking that is not higher than in cat (5 kHz), but likely it is somewhat lower.

Chapter 1

General Introduction

Chapter 1: General Introduction

In the first section, we briefly describe the general context and scientific motivation of this work. In subsequent sections, background information is provided that is necessary for the proper understanding of the chapters of this thesis.

General context and motivation

In our daily lives we are constantly surrounded by sounds. The auditory system has the ability to detect, to analyze and to segregate multiple sound sources. Such ability is very important in our daily lives, for instance, to detect a potential danger out of sight or to have a conversation in a noisy surrounding. Understanding how the auditory system achieves this is not only of fundamental interest, but can help to diagnose hearing disorders, or to improve hearing technology.

The auditory system, from ear to cortex, has a complex pathway, which can be divided into two main parts, the peripheral auditory system and the central auditory system. The peripheral auditory system is the first stage that encodes the sound into a primary neural code which is fed through the auditory nerve to the central auditory system. The central auditory system is the actual neural processing part which re-encodes the primary neural code via extensively interconnected structures (e.g., cochlear nucleus, superior olivary complex, lemniscal nuclei, inferior colliculus, and auditory thalamus and cortex).

In exploring the auditory system, various methods and techniques are employed by researchers in different disciplines. For instance, psychophysicists devise stimulus paradigms to study behavioral; computational neuroscientists build mathematical models to predict the behavior of a system; neurophysiologists study the activity of single neurons; etc. All these techniques share a common goal, namely unraveling the auditory system, but are not all equally suitable to explore the underlying processes in the auditory pathway. Unfortunately, the most incisive techniques to unravel underlying mechanisms, such as the recording of extra-cellular potentials of single neurons, are the most invasive and require intracranial access. These kinds of techniques cannot be routinely applied in humans, who basically can only be studied using indirect techniques, e.g. psychoacoustics and otoacoustic emissions. Therefore, even the most basic properties of the peripheral auditory system (e.g., frequency selectivity and neural phase-locking), which are well characterized in commonly studied laboratory mammals, have only been characterized indirectly in human.

Frequency selectivity and neural phase-locking are the two most fundamental properties of the peripheral auditory nervous system. These properties originate in the cochlea and are pervasive in behavioral and neural responses.

Frequency selectivity refers to the ability to resolve individual spectral components of complex sounds and is important for the accurate perception of e.g. speech. The general assumption of frequency selectivity in human is that it is similar to that in commonly studied laboratory mammals (Glasberg and Moore, 1990; Ruggero and Temchin, 2005). But this view has been challenged by two studies using, indirect, behavioral (Oxenham and Shera, 2003) and otoacoustic (Shera et al., 2002; SFOAEs) methods that reported sharper tuning in humans. That finding was controversial and it has been argued that these results reflect experimental artifacts rather than real species differences (Ruggero and Temchin, 2005; Eustaquio-Martin and Lopez-Poveda, 2011).

Neural phase-locking is the property of auditory nerve (AN) fibers to lock to the instantaneous pressure fluctuations (the fine-structure) in the waveform of sounds. Phase-locking is essential for binaural hearing, i.e. sound localization in azimuth (left-right), and is thought to be invoked in certain aspects of perception, such as pitch, loudness, and speech perception (e.g., Carney, 1994; Moore, 2008). Its role in these perceptual attributes has been discussed for more than a century. Many auditory researchers are convinced that useful phase-locking extends to much higher frequencies than a few kHz (Teich et al., 1993; Heinz et al., 2001; Recio-Spinoso et al., 2005; Moore, 2008). Comparison between unpublished AN fiber data in monkeys from our laboratory (Michelet et al., 2011) and that in cat shows that phase locking in monkey falls-off earlier than in cat. Based on this observation as well as on other arguments (Joris and Verschooten, 2013), we think it is unlikely that phase-locking in humans would extend to higher frequencies than in cat. This debate is not only of academic interest. The issue has become important in the remediation of deafness: recent development in the technology of cochlear implants allows more fine timing control over the electrical stimulation parameters, and the issue regarding the role of temporal firing of action potentials in the normal human auditory nerve is thus of practical importance.

In this work we assessed the two aforementioned fundamental properties of the peripheral auditory system in humans, i.e. frequency selectivity and neural phase-locking. In order to achieve our goal, we developed new stimulus, recording, and analysis techniques to electrophysiologically quantify these processes in humans.

The mammalian peripheral auditory system

The peripheral auditory system converts acoustic energy of sound into neural signals (action potentials). The periphery is traditionally divided into the outer ear, middle ear and the inner ear (Fig. 1).

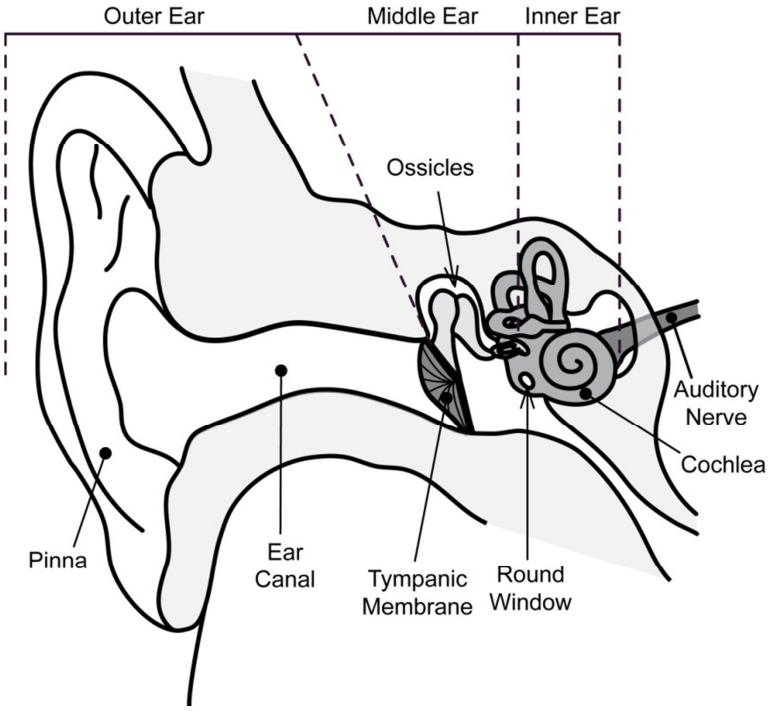


Figure 1. Schematic drawing of the peripheral auditory system. The outer ear guides sound to the tympanic membrane. The middle ear transfers the mechanical vibrations through the ossicles to the cochlea. The inner ear transduces virbrations to electrical signals and performs a spectral analysis of the sound. It provides neural input to the central auditory system via the auditory nerve. Source: redrawn and modified from Russel (2004).

Outer and Middle ear

Sound waves that reach the ear via the pinna (Fig. 1) travel through the ear canal and cause the tympanic membrane to vibrate. These mechanical vibrations are then passed to a chain of ear bones (ossicles: malleus, incus, stapes) suspended in the middle ear. The last bone in the chain, the stapes, is connected to an opening of the cochlea (oval window) via a membrane, so that the mechanical energy conducted through the middle ear is conveyed to the fluid in the inner ear. The tympanic membrane and the bones do not only transfer the vibrational energy, but also act as an impedance matching device between the pressure waves from air and these generated in the fluid of the inner ear (e.g. Peake and Rosowski, 1991). In the middle ear there are also two tiny muscles that are connected to the chain of ossicles. The first one is the tensor tympani which is attached to the malleus and the second is the stapedius muscles which is attached to the stapes. The main function of these muscles is to attenuate (~20 dB) the transduction of vibrational energy to the inner ear by pulling on the ossicles. For

example the muscles contract when the inner ear is exposed to high-intensity sounds (>85 dB SPL); this effect is called the acoustic reflex. Note that the total transfer function is influenced by the acoustical (outer ear) and mechanical auditory periphery (middle ear). The pinna plays a role in the directional hearing, particularly in elevation (up-down) (Blauert, 1983; Voss et al., 1996).

Inner ear

The cochlea is the part where vibrational energy, passed by the stapes to the fluid, is transduced into electrical signals. The cochlea is a snail-shaped hollow chamber encased in bone. It is divided into three fluid filled chambers separated by a membrane, the scala vestibuli, the scala tympani and the scala media (Fig. 2). The scala tympani and the scala vestibuli are filled with perilymph (low K^+ concentration), and merges at the helicotrema which is located at the apex of the cochlea.

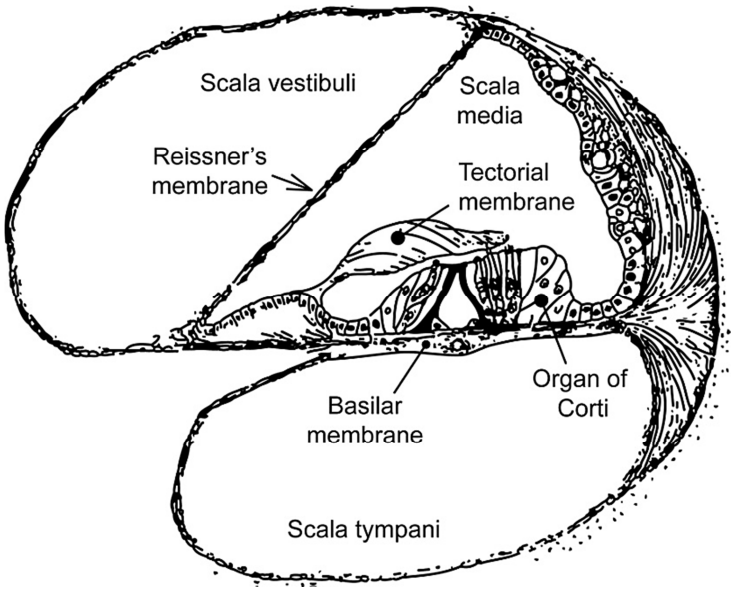


Figure 2. Schematic diagram of a cross section of the cochlea. The illustration shows the three major fluid-filled chambers, called scalae, which are separated by the two membranes, the basilar membrane and the Reissner's membrane. The organ of Corti lies on top of the basilar membrane. Modified, from Rasmussen (1943).

The scala media is located between the scala vestibuli and scala tympani and is formed by two membranes, Reissner's membrane and the basilar membrane (BM), respectively. The scala media is filled with endolymph (high K^+ concentration) and has a positive potential (80-120 mV) with respect to the other scalae. At the base of the cochlea there are two openings (windows) in the bone covered by a membrane, the oval window and the round window. The oval window is located in the scala vestibuli and accommodates the stapes, as mentioned

earlier. The other window, the round window (RW; Fig. 1), is located in the scala tympani and allows movement of the fluid within the cochlea Thus, every time the oval window moves inward, the round window moves outward.

Mechanical response of the basilar membrane

The cross-section of the scalae and the mechanical properties of the cochlea changes systematically along its length. From base to apex, the area of the cross section become gradually smaller and the width of the BM become wider; the BM is up to 10 times larger at the apex compared to the base. Because of these gradual changes, the mechanical resonance of the BM is frequency dependent and is tuned along its length. High frequency tones cause maximal vibration near the base of the cochlea and low frequency tones near the apex of the cochlea (Fig. 3A; from cadaver cochleae). In the mammalian cochlea, the energy causing the BM to vibrate is mainly supplied by a relative slow traveling wave that propagates from base to apex (von Békésy, 1960; Faingold et al., 1991; Robles and Ruggero, 2001). (Fig. 3B)

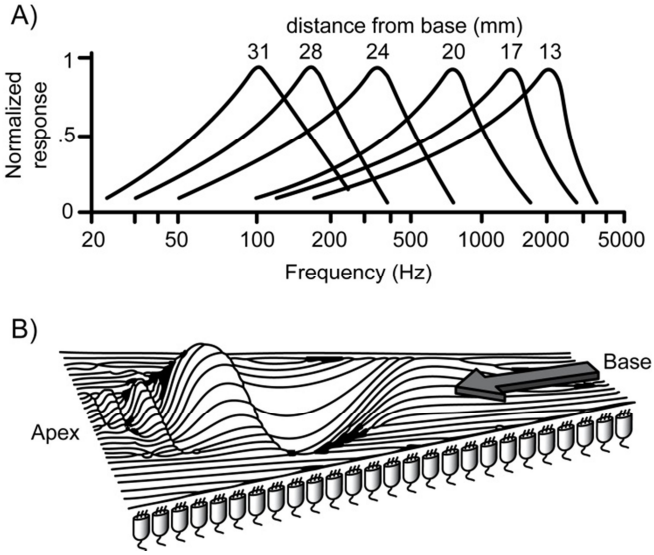


Figure 3. A) Illustration of the response of six places along the length of the cochlea to tones at different frequencies. Results are from cadaver cochleas. Redrawn and modified from von Békésy and Rosenblith (1951).

B) Snapshot illustration of the uncoiled basilar membrane in response to a tonal stimulus. This stationary picture occurs only at one point in time. The arrow indicates the propagation direction of the traveling wave, from base to apex.

Sensory receptor

The organ of Corti (Fig. 4) is the sensory receptor that converts the mechanical vibrations of the BM into electrical signals. It is located in the scala media, on top of the BM, and is present along the length of the BM. It consists of structural supporting cells that imbed the actual sensory cells and their innervating AN fibers. There are two types of sensory cells, the inner hair cells (IHCs) and outer hair cells (OHCs). The hair cells are aligned along the BM; there are 3 or 4 OHCs for every IHC (~1500 IHCs). The IHCs are completely

surrounded by supporting cells, whereas the OHCs are only supported at the top and the bottom. Each IHC is innervated by ~20 afferent fibers (i.e. conveying information to the central nervous system) from single spiral ganglion cells which provide the majority (~95%) of the neural inputs through the AN to the brainstem. The outer hair cells are thought to be primarily involved in the enhancement of the mechanical vibrations, and are the hypothesized substrate of the "cochlear amplifier". The OHC are innervated by efferent fibers, which allow control of the central nervous system over the cochlea. Efferent activation causes a decrease in gain of the cochlear amplifier.

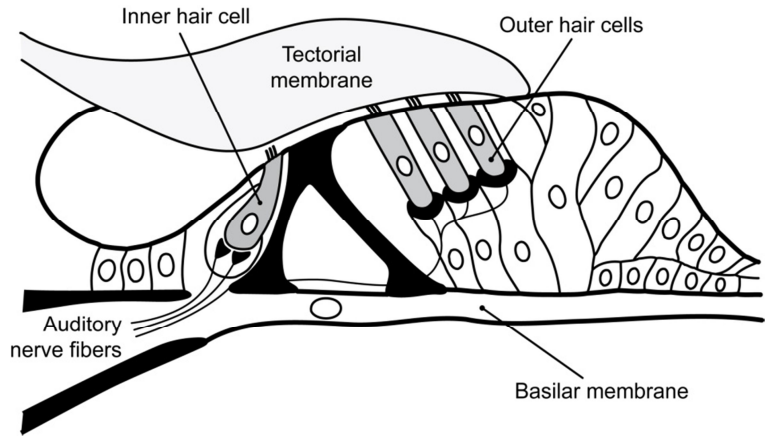


Figure 4. Illustration of the organ of Corti. The inner and outer hair cells are indicated in gray. The gelatinous tectorial membrane contacts the stereocilia of the OHCs and overlies the stereocilia of the IHCs. Source: redrawn and modified from Yates et al. (1992).

Mechano-electrical transduction

The hair cells transduce the mechanical vibrations of the BM into a receptor potential (transmembrane voltage) by means of stereocilia that are located on top of the hair cells. When stereocilia are deflected, mechanically gated ion channels are opened and positive charged ions (K^+) enter the cell (transduction current) and change (depolarize) the receptor potential of the cell. The depolarization of the receptor potential results in the activation of voltage gated calcium channels. Calcium ions enter the cell and trigger the release of neurotransmitters at the basal end of the cell, which diffuse across the synaptic cleft and induce action potentials in the nerve terminal of the spiral ganglion neurons.

Cochlear amplifier and compression

The extraordinary sensitivity and the high frequency tuning of the mammalian cochlea (Robles and Ruggero, 2001) are not simply caused by the passive mechanics of the inner ear but are the result of what is called the "cochlear amplifier" (Davis, 1983). The detail of this enhancement is not completely understood, but it is generally accepted that outer hair cells are actively involved in this amplification process. Two candidate mechanisms have been proposed for this active involvement: stereociliary motility (Crawford and Fettiplace, 1985) and somatic motility (Brownell et al., 1985). The first mechanism was demonstrated in turtle but its role in mammals is debated. Somatic motility (motor protein prestin) is required for amplification in the mammalian cochlea (Liberman et al., 2002; Dallos et al., 2008). This means that OHCs perform not only a mechano-electrical transduction like the IHCs, but also an electromechanical transduction, whereby the length of the OHC changes (*in vitro*) as a function of receptor potential. It is believed that this electromotility injects mechanical energy to the cochlear partition on a cycle-by-cycle basis (Dallos and Evans, 1995).

The amplification is level dependent (Fig. 5): at low sound levels the amplification is highest, at middle sound levels it decreases and at high levels it is negligible. Often one refers to this level dependent effect as cochlear compression. As a side effect, the non-linear response of the BM causes complex interactions between spectral components in the stimulus which create additional intermodulation distortion products.

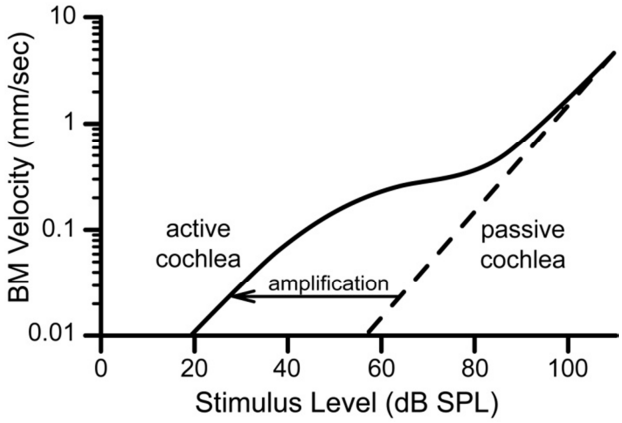


Figure 5. Illustration of a typical basilar membrane input-output function with and without amplification. Source: redrawn and modified from Chadwick, (1998).

The active process of the OHCs is associated with the phenomenon of otoacoustic emission. Otoacoustic emission is the generation of vibration in the inner ear which is transmitted back in the reverse direction through the middle ear and to the ear canal where it can be measured acoustically. These emissions can be spontaneous or can be evoked by external sound stimulation. External otoacoustic emission (Kemp, 1978) are often applied as a

research tool, e.g. to estimate frequency tuning (Shera and Guinan, 1999) or as an objective diagnostic tool by audiologists.

Neural responses

The main characteristics (frequency tuning, sensitivity, and non-linearity) observed in the mechanical responses of the BM (Robles and Ruggero, 2001) are reflected in the responses of auditory nerve fibers as a rate and temporal code. Both codes, the (average) discharge rate and discharge time, relay information regarding stimulus level and frequency. The following two sections describe the two fundamental properties of the responses of single AN fibers, neural frequency tuning and neural phase-locking. These properties are typically studied with pure tones because of their simple interpretation. Note that AN fibers exhibit, besides the former fundamental properties, also other response properties such as a spontaneous rate, a threshold level, a dependence of rate versus level, a restricted dynamic range, saturation in rate, etc.

Neural Frequency selectivity

AN fibers that innervate IHCs along the BM are most sensitive to a particular "characteristic" tonal frequency. The tuning of AN fibers is classically described by a frequency-threshold tuning curve (FTC; Fig. 6). This curve represents the tonal level as a function of frequency needed to satisfy a threshold response criterion (i.e. a certain level of neural activation, e.g. a fixed number of action potentials).

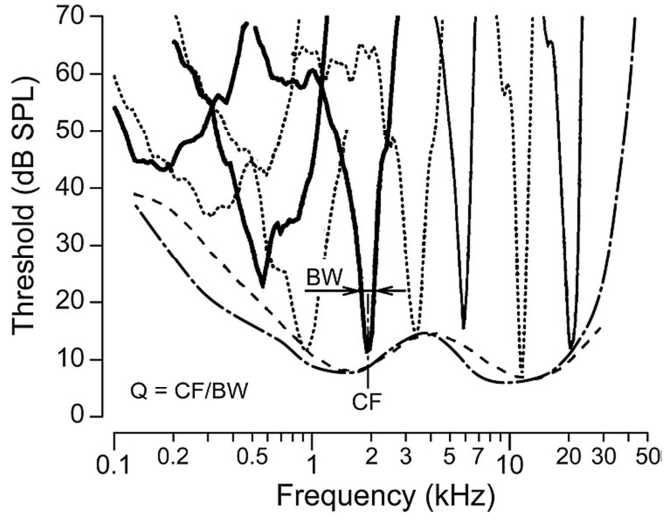


Figure 6. Neural frequency-threshold tuning curves in monkey (macaque) for 9 selected frequencies. The tips of the most sensitive tuning curves indicated by the black dashed line are in reasonable agreement with the behavioral audiogram (dashed dotted line). The sharpness is often indicated by a Q-factor. Figure modified from Joris et al. (2011).

The FTC typically exhibits a sharp V-shape and is similar across species, but the frequency range and sharpness of tuning varies from species to species. The frequency corresponding to the most sensitive level, here referred to as the tip or nadir of the FTC, is called the characteristic frequency (CF). Sharpness of frequency tuning is commonly expressed as a relative measure, the Q-factor. For single AN fibers, the Q_{10} factor is traditionally used. The Q_{10} is calculated as the reciprocal of the relative bandwidth at 10 dB above the nadir, i.e. $Q_{10} = CF/BW_{10}$. The sharpness of tuning of single AN fibers depends on CF (Figs. 6,7). Fibers with high CFs have in general sharper tuning than those with low CFs.

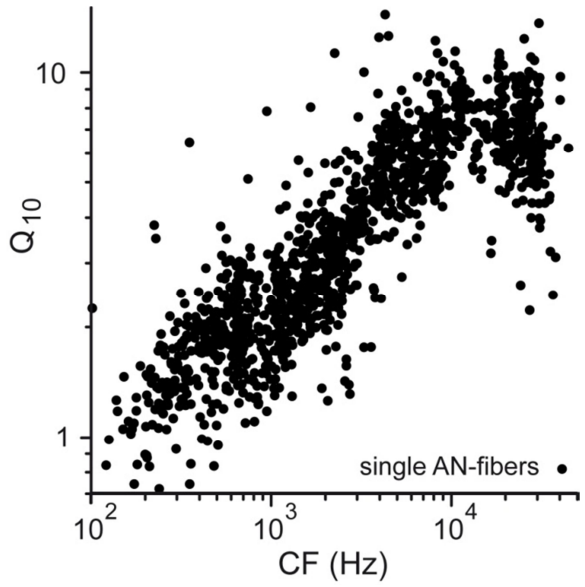


Figure 7. Sharpness of tuning of single auditory nerve fibers as a function of frequency in cat. Sharpness of tuning is expressed as a Q_{10} factor. Source: data from the database of our laboratory.

Phase-locking

The other fundamental property we discuss in this introduction is phase-locking, which is the synchronization of AN activity to the stimulus waveform. This property arises from the receptor potential (membrane voltage) of IHCs which has the ability to follow the fine structure in the waveform (Palmer and Russell, 1986). Basically the receptor potential can be divided in to two components, a level dependent DC component and AC component, which is waveform dependent (Fig. 8b, left panel). For low stimulus frequencies, the ratio between these components (AC/DC) is fairly constant, but decreases rapidly above a certain cut-off frequency so that ultimately only the DC component remains (Fig. 8b, right panel). The low-pass filter process involved in this attenuation is thought to originate from the time constant of the membrane (Palmer and Russell, 1986; Russell and Palmer, 1986), but due to

the large roll-off (≥40 dB/decade) other filtering mechanisms are likely involved in the loss of synchronization (Kidd and Weiss, 1990).

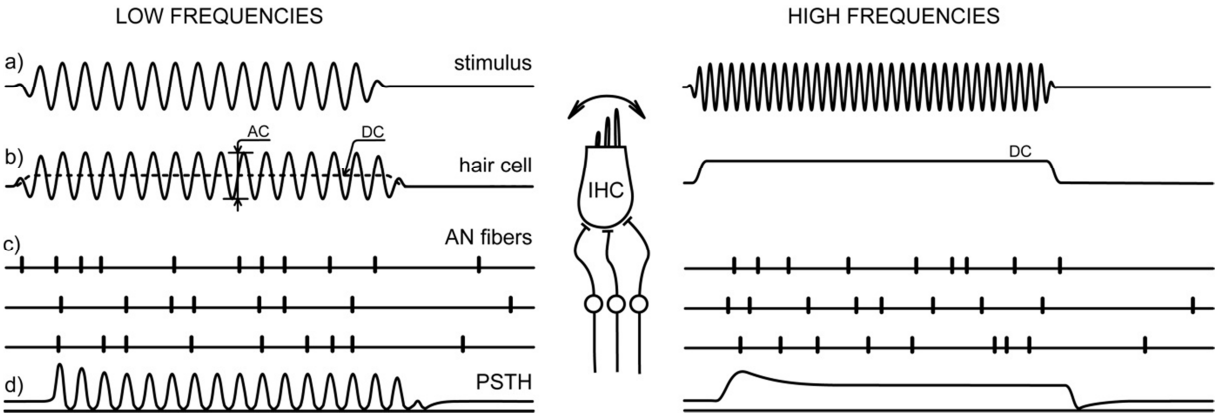


Figure 8. Illustration of peripheral responses to a low (left panel) and high (right panel) frequency tone. a) Acoustic stimulus. b) Receptor potential of the IHC. c) Three illustrations of evoked time events in different nerve fibers. d) Illustration of a post-stimulus time histograms of the response of a nerve fiber. Source: modified from Joris (2010).

AN fibers that innervate IHCs show synchronized discharge activity to the AC potential of their IHCs. The fibers typically do not fire on every cycle of a low frequency tone, (Fig. 8c, left panel), but when they do, they fire at a preferred position or phase aligned in time with the (tonal) stimulus (Tasaki, 1954; Kiang et al., 1965). This is illustrated in Fig. 8d (left panel) by the post-stimulus time histogram (PSTH), which depicts the probability of discharge in the AN to the repeated stimulus. Note that the AC component of the probability of discharging to a low frequency pure tone is non-linear and shows half wave rectification. For high frequency tones (right panel), the responses and PSTHs (Fig. 8c,d) show only an increased probability of firing during stimulation; the activity synchronized to the stimulus waveform, which is present for low frequency tones, is gone. There is another temporal component of importance for the work presented here: enhanced and reduced activity at the on- and off-set of the stimulus (cf. Fig. 8d). This is visible in the PSTH as a peak at stimulus onset and as a dip at stimulus offset. These features are due to adaptation, presumably caused by enhanced or reduced availability of IHC neurotransmitter. Remark the non-zero probability of discharging during no stimulation; this is due to the spontaneous rate, which differs between fibers.

Neural phase-locking in AN fibers is classically expressed as the vector strength (VS), which is a kind of synchronization index that quantifies phase-locking with a single number between zero and one. Random spiking means zero vector strength (cf. Fig. 8c,d) and a vector strength of one means perfect synchronization whereby all spikes occur at exactly the same phase during the period of a tone (Goldberg and Brown, 1969). The vector strength is the normalized vector sum of unit spike events represented on a stimulus-phase polar plot and can be calculated as

$$VS = \frac{1}{n}|\nu| \qquad (1)$$
$$\text{and} \quad \nu = \sum_{i=1}^{n} e^{j\varphi_i}$$

with φ_i the (wrapped) stimulus phase at spike event i and n the number of spikes. The preferred angle of the VS is then calculated as $\theta = arg(\nu)$. It is common to visualize phase-locking with a period or cycle histogram (Fig. 9B), where the unit spike events are collected in time bins. Note that the vector strength can also be obtained from the spectrum of the cycle histogram as the first harmonic divided by the DC value.

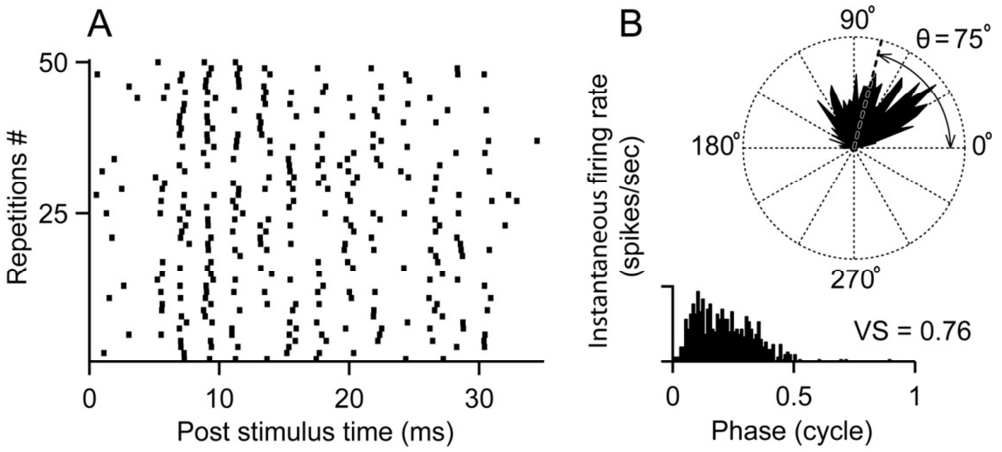


Figure 9. Response of an AN fiber to 50 stimulus presentations of a 30-ms tone at 220 Hz. A) Dot raster: every dot represents a single spike event in response to a pure tone, repeated in 50 trials. Each horizontal series of dots is from another trial. B) Period or cycle histogram of the response in A, represented in a polar and Cartesian coordinate system. The VS is here 0.76 with a preferred angle of 75°. Source: Joris PX.

Similar to the ratio of the AC and DC potentials in IHCs is also the maximum VS frequency dependent (Fig. 10). At low frequencies the VS is fairly constant, but declines rapidly from a certain cutoff frequency until it is no longer significant anymore above a certain upper frequency limit. The cutoff and the frequency limit of phase-locking are species dependent (Weiss and Rose, 1988a), but are unknown in human.

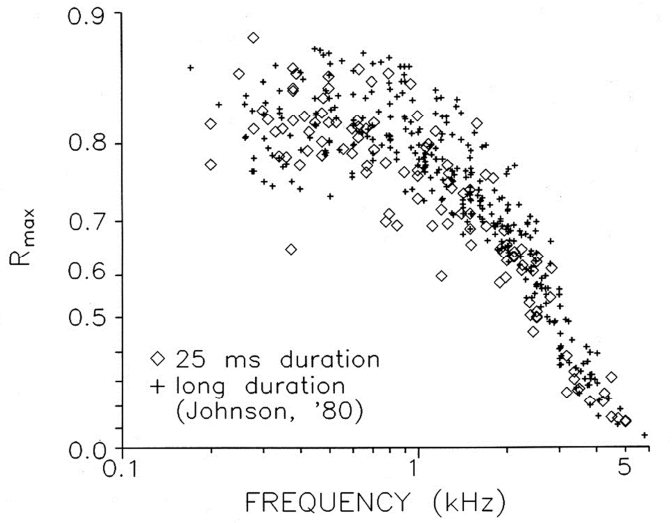


Figure 10. Maximum vector strength (Rmax) in cat as a function of CF. Source: Joris et al. (1994b).

Forward masking

Masking is the reduction in the response (or perception in case of psychophysics) to a signal (the "probe") by the presence of another signal (the "masker"). Often the probe is a pure tone and the masker is a broadband or bandlimited noise, and the behavioral task is to detect the tone. There are two types of masking; spectral masking (simultaneous masking) where the signal of the masker is simultaneously present with the signal (tone) of the probe; and temporal masking (non-simultaneous masking) where the probe is preceded (forward masking) or followed (backward masking) by a masker. Forward masking is the preferred method in this study, because backward masking is not well understood and is thought to be a "higher level" phenomenon, and simultaneous masking introduces nonlinear cochlear effects like two-tone suppression and can generate inter-modulation products.

Forward masking is thought to reflect adaptation of the sensory system, and is typically modeled based on how the firing rate of AN fibers changes over time in response to a stimulus. There are at least two distinct processes involved in the adaptation of AN fibers (Westerman and Smith, 1984): a process with a short (fast) time constant (Fig. 11A, τ_f) of a few milliseconds; and a concurrent process with a larger (slow) time constant (Fig. 11A, τ_s) of about a few tens of milliseconds. These adapting processes are not only active during stimulation but also thereafter, when the system re-adapts to its pre-stimulus condition. This re-adaptation is visible in a PSTH by the recovery of the spontaneous activity (cf. Fig. 11A, τ_r). Now, when a signal (the probe) with typical low intensity is placed in the recovery period of the preceding signal (masker), then the response of this signal (red, Fig. 11B) is reduced; the larger the influence of the masker on the probe (e.g. because of frequency or relative level), the larger the reduction of the probe response (Fig. 11B).

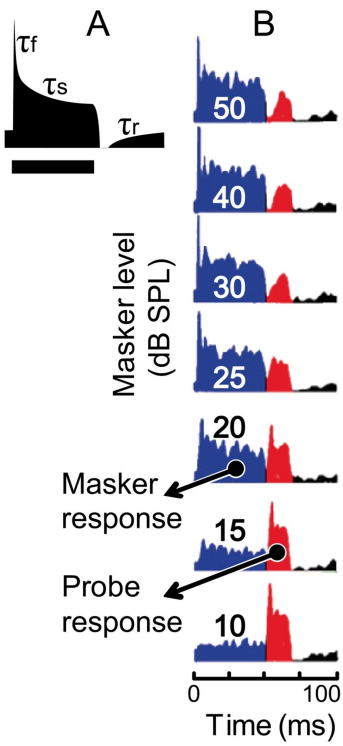


Figure 11.

A) Illustration of adaptation in a PSTH for a typical AN fiber; the stimulus duration is indicated by the bar.

B) Forward masking in an AN fiber for different masker levels. Blue: masker response; red: probe response. Masker: tone 6.5 kHz, 50 ms, SPL as indicated in PSTH. Probe: tone 6.5 kHz, 22 dB SPL, 15 ms. Masker-probe interval 2ms. Figure modified from Harris and Dallos (1979).

In this study we use forward masking as a tool to selectively "deplete" a portion of the AN fibers, either to assess frequency selectivity (Chapter 2 and 3) or to disambiguate neural activity from hair cell activity (neural phase-locking, Chapter 4 and 5). The principle to assess frequency selectivity by forward masking is illustrated for a single AN fiber in Fig. 12 and is explained in its caption.

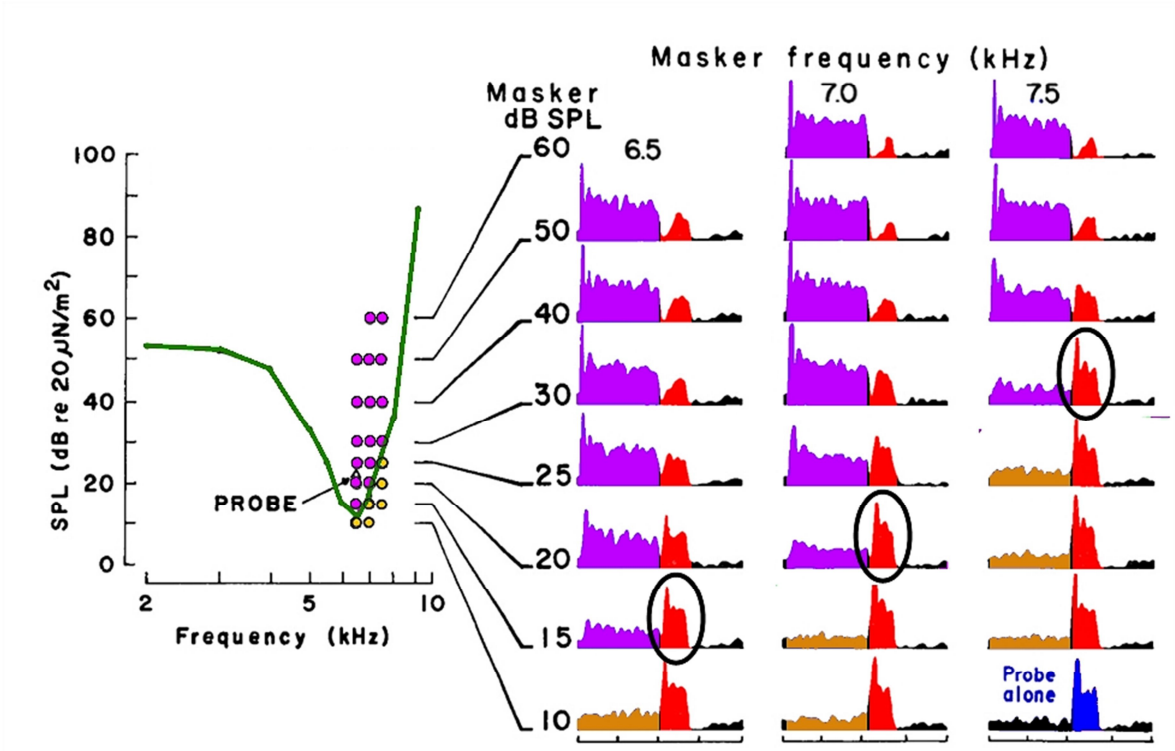


Figure 12. Forward masking in relation to the excitatory response area of a single AN fiber (Fig. 11). A 50 ms duration masker is varied in intensity (10-60 dB SPL; rows) and frequency (6.5, 7.0, 7.5 kHz; columns) as indicated by the matrix of circles on the FTC. The green line represents the normal FTC curve of the AN fiber. The PSTHs on the right hand side illustrate how probe response magnitude (red colored areas) changes as a function of masker level and masker frequency. The circles and masker responses with the purple color indicate the masker conditions that reduce the probe response with 20% (circled) or more with respect to the control (probe alone, PSTH in right-bottom corner). Note that the encircled probe responses with the same 20% reduction have similar masker frequency/levels than those of the FTC. The curve drawn through these masker frequency/levels is called a masking tuning curve and is a good relative approximation of the FTC. Probe stimulus: tone at CF, level: 10 dB above threshold, 15 ms duration. Figure modified from Harris and Dallos (1979).

Cochlear potentials

The ultimate goal of this study is to derive the two fundamental properties, discussed above, in human. Unfortunately, as explained earlier in section A, it is not possible to measure single AN responses in volunteers. The best alternative for these responses are externally measurable far-field potentials.

The cochlea has a number of internal DC and AC (mass) potentials which reflect cochlear activity. Three of these potentials, namely the summating potential, the cochlear microphonic (CM) and the whole nerve potential, are stimulus related and are measurable as a compound potential outside the cochlea. The most interesting potential for this study is the whole nerve potential, because it reflects compound activity of the auditory nerve. The whole nerve potential can be subdivided in two potentials: the compound action potential (CAP) and the auditory nerve neurophonic (NP), which reflect different features of the neural response. We use these two potentials to study frequency selectivity and phase-locking.

One of the challenges when using these two mass potentials is to separate them from underlying receptor potentials. To achieve this, different techniques can be applied that make use of spectral and temporal properties of the potentials. In what follows we briefly discuss the three AC potentials of interest (CM, CAP and NP) and indicate their most important properties. In this study we do not make use of the summating potential (DC component), it is always rejected in the recordings by a high-pass filter (normally set to 100 Hz). Since this component is not used in this study, it will not be further discussed.

Cochlear Microphonic (CM)

The cochlear microphonic comes from the receptor cells and is believed to be primarily generated by the OHCs. It is the alternating potential associated to the transduction currents flowing through the stereocilia of the many receptor cells. This potential is a fairly good replication of the time varying sound and extends to the highest frequencies. For moderated and low intensity tones the CM is almost perfectly linear.

Compound Action Potential (CAP)

The CAP is an AC potential that appears in the far-field as a large negative peak followed by a positive peak and often followed by other several decaying negative and positive peaks. The presence of the CAP is associated with rapid temporal changes in the stimulus, such as clicks, on-sets and off-sets. The negative peaks in the CAP reflect the summed synchronized discharge of a large neural population in the AN (Fig. 13) (Goldstein and Kiang, 1958). The CAP has some specific and inherited properties from the underlying responses of the AN fibers. a) The CAP-waveform is restricted in time and is rather independent of stimulus polarity, e.g. for tone pips alternating in phase by π. b) The magnitude of the CAP increases monotonically with stimulus intensity, whereas the underlying number of extracellular action potentials of AN fibers does not or not in the same

extent. c) The position of the maximum negative deflection is level dependent and advances in time with increasing stimulus level. d) CAPs are, like the underlying action potentials, susceptible to adaptation and can be reduced by a forward masker. The latter property is extensively used in this study.

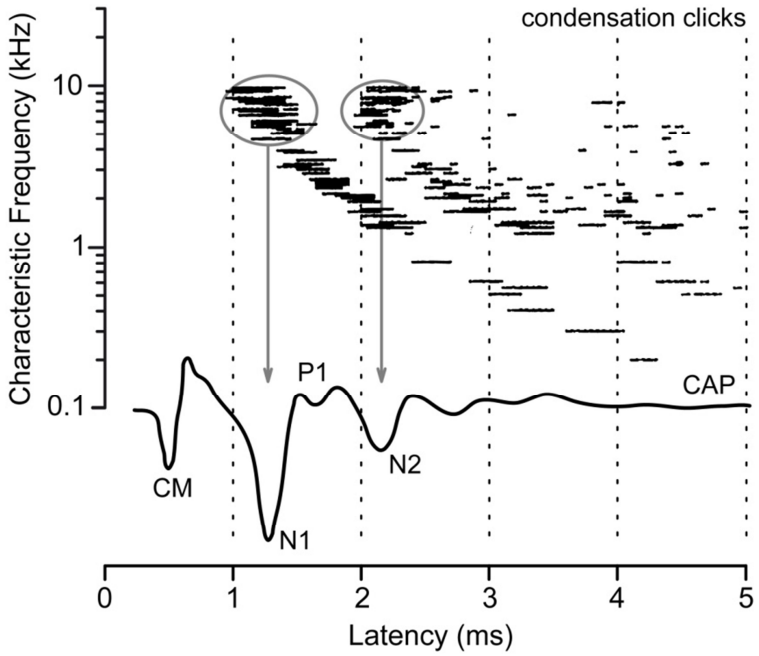


Figure 13. Times of AN fiber responses to condensation clicks compared with potentials recorded near the round window. Stimuli: 10 clicks/s, 88 dB SPL. The lines in the upper panel indicate the times of firing of action potentials in AN fibers of different CF (ordinate). CM: Cochlear micro-phonic potential; N1: First negative deflections of neural activity. Source: redrawn and modified from Kiang (1984).

Auditory Nerve Neurophonic (NP)

The NP is an AC potential from the gross synchronized activity of the AN and was first noticed by Wever and Bray (1930) when recording potentials from the AN in cat. The NP is, like the CAP, a component of the whole nerve potential, but is distinguished from the CAP because it is an ongoing potential which has the ability to follow the fine structure of the stimulus up to a certain frequency. The NP and CM are easily and often confused because they are both synchronized to the stimulus and are always simultaneously present to some degree.

The NP is an interesting potential because it reflects neural phase-locking in AN fibers and is therefore a potential candidate to predict the frequency limit of neural phase-locking in human. Moreover, the NP is not only detectable at the AN itself but has been reported to be present in the compound potential measured at or near the cochlear RW in animals (Henry, 1995). In this work we confirm this observation and show that this is also true in humans. Because this location is accessible in human, it allowed us to make recordings of neural phase-locking in volunteers with normal hearing. Unfortunately, at this location the NP and

CM are strongly entangled and in addition the CM extends to higher frequencies than the strong declining NP. Therefore, for the determination of the limit of neural phase-locking it is of utmost importance to eliminate the CM. This is one of the challenges in this study and the subject of Chapter 4. Without going into too much detail at this point, there are two important properties that distinguish the NP from the CM. First the NP is like the responses of AN fibers adaptable and can be masked by a forward masker (section *forward masking*), while the CM is not of neural origin and not forward maskable (Snyder and Schreiner, 1985). Second, because of the rectifying effect of the IHCs (remind that action potentials are only triggered during the depolarization phase, not the hyperpolarization phase), the NP is non-linear and contains a large 2nd harmonic. In contrast, the CM is highly linear over a wide intensity range. Therefore, the 2nd harmonic is often used as an indicator for the presences of the neural contribution. Next, in additional to these two properties there is a feature that we will use as argument for the cancellation of CM. Namely, a short latency between NP and CM because of several delays like a synaptic delay and axonal propagation delay.

Objectives

Our ultimate goal is to electrophysiologically quantify neural frequency selectivity and stimulus related synchronization (neural phase-locking) in human. In laboratory animals, these properties are typically determined from extracellular responses of single AN fibers. Unfortunately, this requires invasive recording techniques with intracranial access and penetrating electrodes into the auditory nerve. Therefore, it cannot be applied in normal hearing humans, so that indirect techniques necessarily have to be used in an attempt to quantify these properties. The results of the different techniques applied so far (psychophysics and otoacoustic emissions) are very indirect, often vary widely, and their interpretation is often controversial so that there is no general agreement on the quantitative limits of frequency selectivity and phase-locking in humans. This brings us to two specific questions that have recently received much attention amongst researchers in the auditory field. First, do humans have sharper frequency tuning than commonly studied animals and second, do humans have a higher upper frequency limit of phase-locking than commonly studied animals. To answer these two questions, we developed for each question a method that allows us to obtain a measure for the two properties, and compare the outcome to that from single AN fibers for different species. Our methods make use of far-field potentials. Before making recordings in human we developed, parameterized, optimized and verified stimulus and analysis paradigms in cat, chinchilla and monkey (macaque). As a final step, we then

developed a new recording technique to obtain the first electrophysiological data reflecting peripheral neural activity in normal hearing volunteers. The questions and methods led us to several studies that are the subject of the next four chapters.

Study 1 – In a first study we developed a method in cat and chinchilla to estimate neural frequency selectivity from whole nerve potentials. More specific, CAPs were used in combination with a paradigm similar to what was used in a psychophysical study reporting exceptional sharp tuning in humans. In addition, a parametric study was performed to assess the influence of level. Finally, sharpness of tuning was compared with that in individual AN fibers.

Study 2 – In a second study we developed a procedure to measure cochlear potentials from the cochlear promontory or niche of the round window through the middle ear in volunteers with normal hearing. The same paradigm as that of study 1 was used. CAP-tuning was also measured in monkeys. The results amongst species were compared and used to estimate neural frequency selectivity in human.

Study 3 – In the next study a method was developed to isolate the neural phase-locked contribution (neurophonic) present in the cochlear potential. The method was assessed at the cochlear round window in cat. Moreover, the method was verified using an effective neural blocker (Tetrodotoxin). The upper limit of phase-locking was determined and compared to that of individual AN fibers.

Study 4 – In the final study cochlear potentials from the cochlear promontory or niche of the round window were recorded in monkey (macaque) and in human volunteers using the paradigm developed in the study 3. Neural phase-locking was quantified and the upper frequency limit of phase-locking was determined. The results of cat, monkey and human are compared and discussed.

REFERENCES

Blauert J (1983) Spatial Hearing. Cambridge: MIT Press.

Brownell WE, Bader CR, Bertrand D, de Ribaupierre Y (1985) Evoked mechanical responses of isolated cochlear outer hair cells. Science 227:194-196.

Carney LHC (1994) Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness. HearRes 76:31-44.

Chadwick RS (1998) Compression, gain, and nonlinear distortion in an active cochlear model with subpartitions. Proceedings of the National Academy of Sciences 95:14594-14599.

Crawford AC, Fettiplace R (1985) The mechanical properties of ciliary bundles of turtle cochlear hair cells. The Journal of physiology 364:359-379.

Dallos P, Evans BN (1995) High-frequency motility of outer hair cells and the cochlear amplifier. Science 267:2006-2009.

Dallos P, Wu X, Cheatham MA, Gao J, Zheng J, Anderson CT, Jia S, Wang X, Cheng WH, Sengupta S, He DZ, Zuo J (2008) Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. Neuron 58:333-339.

Davis H (1983) An active process in cochlear mechanics. Hear Res 9:79-90.

Eustaquio-Martin A, Lopez-Poveda EA (2011) Isoresponse versus isoinput estimates of cochlear filter tuning. J Assoc Res Otolaryngol 12:281-299.

Faingold CL, Boersma A, Caspary DM (1991) Involvement of GABA in acoustically-evoked inhibition in inferior colliculus neurons. Hear Res 52:201-216.

Glasberg BR, Moore BC (1990) Derivation of auditory filter shapes from notched-noise data. Hearing research 47:103-138.

Goldberg JM, Brown PB (1969) Response of Binaural Neurons of Dog Superior Olivary Complex to Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. J Neurophysiol 22:613-636.

Goldstein MH, Kiang NYS (1958) Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. The Journal of the Acoustical Society of America 30:107-114.

Heinz MG, Colburn HS, Carney LH (2001) Evaluating auditory performance limits: i. one-parameter discrimination using a computational model for the auditory nerve. Neural Comput 13:2273-2316.

Henry KR (1995) Auditory nerve neurophonic recorded from the round window of the Mongolian gerbil. Hearing research 90:176-184.

Joris PX (2010) Het snelle slakkenhuis. In: Lessen voor de eenentwintigste eeuw – In gesprek met morgen (Raymaekers B, ed), pp 173-192. Leuven: Universitaire Pers Leuven.

Joris PX, Verschooten E (2013) On the limit of neural phase-locking to fine-structure in humans. Basic Aspects of Hearing: Physiology and Perception Editors: BCJ Moore, RD Patterson, I Winter, RP Carlyon, and HE Gockel (Springer, NY), in press.

Joris PX, Carney LH, Smith PH, Yin TC (1994b) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. Journal of neurophysiology 71:1022-1036.

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proc Natl Acad Sci U S A 108:17516-17520.

Kemp DT (1978) Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 64:1386-1391.

Kiang NYS (1984) Peripheral neural processing of auditory information. In: Handbook of Physiology. Section 1, Volume III, part Edition (Brookhart JM, Mountcastle VB, eds), pp 639-674. Bethesda: American Physiological Society.

Kiang NYS, Watanabe T, Thomas EC, Clark LF (1965) Discharge patterns of single fibers in the cat's auditory nerve, 35 Edition. MIT Press, Cambridge: Research Monograph No 35.

Kidd RC, Weiss TF (1990) Mechanisms that degrade timing information in the cochlea. Hear Res 49:181-207.

Liberman MC, Gao J, He DZ, Wu X, Jia S, Zuo J (2002) Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. Nature 419:300-304.

Michelet P, Mc Laughlin M, van der Heijden M, Joris P (2011) Synchronization to pure and amplitude-modulated tones in the auditory nerve of macaque monkey. Assoc Res Otolaryngol Abstract 34:223.

Moore BC (2008) The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. Journal of the Association for Research in Otolaryngology : JARO 9:399-406.

Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J Assoc Res Otolaryngol 4:541-554.

Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair cells. Hear Res 24:1-15.

Peake WT, Rosowski JJ (1991) Impedance matching, optimum velocity, and ideal middle ears. Hear Res 53:1-6.

Rasmussen AT (1943) Outlines of neuro-anatomy. Dubuque (Iowa): Wm. C. Brown.

Recio-Spinoso A, Temchin AN, van Dijk P, Fan YH, Ruggero MA (2005) Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. Journal of neurophysiology 93:3615-3634.

Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. Physiological reviews 81:1305-1352.

Ruggero MA, Temchin AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. Proceedings of the National Academy of Sciences of the United States of America 102:18614-18619.

Russel AD (2004) Introduction to Psychology: Wadsworth Publishing.

Russell I, Palmer A (1986) Filtering due to the inner hair-cell membrane properties and its relation to the phase-locking limit in cochlear nerve fibres. In: Auditory Frequency Selectivity (Moore BCJ, Patterson RD, eds), pp 199-207. New York: Plenum Press.

Shera CA, Guinan JJ, Jr. (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am 105:782-798.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proceedings of the National Academy of Sciences of the United States of America 99:3318-3323.

Snyder RL, Schreiner CE (1985) Forward masking of the auditory nerve neurophonic (ANN) and the frequency following response (FFR). Hearing research 20:45-62.

Tasaki I (1954) Nerve impulses in individual auditory nerve fibers of guinea pig. J Neurophysiol 17:97-122.

Teich M, Khanna S, Guiney P (1993) Spectral characteristics and synchrony in primary auditory-nerve fibers in response to pure-tone acoustic stimuli. J Stat Phys 70:257-279.

von Békésy G (1960) Experiments in hearing. New York: Mc Graw Hill.

von Békésy G, Rosenblith WA (1951) In S. S. Stevens (Ed.). The mechanical properties of the ear. New York: Wiley.

Voss SE, Rosowski JJ, Peake WT (1996) Is the pressure difference between the oval and round windows the effective acoustic stimulus for the cochlea? J Acoust Soc Am 100:1602-1616.

Weiss TF, Rose C (1988a) A comparison of synchronization filters in different auditory receptor organs. Hear Res 33:175-180.

Westerman LA, Smith RL (1984) Rapid and short-term adaptation in auditory nerve responses. Hear Res 15:249-260.

Wever EG, Bray CW (1930) Action Currents in the Auditory Nerve in Response to Acoustical Stimulation. Proc Natl Acad Sci U S A 16:344-350.

Yates GK, Johnstone BM, Patuzzi RB, Robertson D (1992) Mechanical preprocessing in the mammalian cochlea. Trends in neurosciences 15:57-61.

Chapter 2

Auditory nerve frequency tuning measured with forward-masked compound action potentials

Verschooten E, Robles L, Kovačić D and Joris PX (2012)
J Assoc Res Otolaryngol. Dec;13(6):799-817

ABSTRACT

Frequency selectivity is a fundamental cochlear property. Recent studies using otoacoustic emissions and psychophysical forward-masking suggest that frequency selectivity is sharper in human than in common laboratory species. This has been disputed based on reports using compound action potentials (CAPs), which reflect activity in the auditory nerve and can be measured in humans. Comparative data of CAPs, obtained with a variety of simultaneous masking protocols, have been interpreted to indicate similarity of frequency tuning across mammals and even birds. Unfortunately, there are several issues with the available CAP measurements which hamper a straightforward comparison across species. We investigate sharpness of CAP tuning in cat and chinchilla using a forward masking notched-noise paradigm – which is less confounded by cochlear nonlinearities than simultaneous masking paradigms and similar to what was used in the psychophysical study reporting sharper tuning in humans. Our parametric study, using different probe frequencies and notch widths, shows relationships consistent with those of auditory-nerve fibers (ANFs). The sharpness of tuning, quantified by Q_{10} factors, is negatively correlated with probe level and increases with probe frequency, but the Q values are generally lower than the average trend for ANFs. Like the single fiber data, tuning for CAPs is sharper in cat than in chinchilla, but the two species are similar in the dependence of tuning on probe frequency and in the relationship between tuning in ANFs and CAP. Growth-of-maskability (GMB) functions show slopes <1 indicating that with increasing probe level the probe is more susceptible to cochlear compression than the masker. The results support the use of forward-masked CAPs as an alternative measure to estimate ANF tuning and to compare frequency tuning across species.

INTRODUCTION

Frequency tuning is a fundamental property of responses from the cochlea and higher level structures in the auditory system. Frequency tuning of the cochlea and auditory-nerve fibers (ANFs) is most commonly studied by determining the stimulus intensity needed to achieve a criterion response (e.g., basilar-membrane displacement or spike-rate) over a range of pure-tone frequencies. The resulting typically V-shaped curve is referred to as a frequency threshold curve (FTC), and from it the characteristic frequency (CF) and a measure for sharpness of tuning are obtained. By sharpness of tuning we refer to the width of the V-shape relative to the CF, commonly expressed in terms of the so-called quality (Q) factor (see METHODS).

In humans, sharpness of frequency tuning has been studied with a variety of behavioral and physiological methods but is unknown for single ANFs because their study requires invasive recordings. Despite the lack of single ANF measurements, it is generally thought that frequency tuning in humans is similar to that in commonly studied laboratory mammals (Ruggero and Temchin, 2005). On the other hand, an increasing number of studies report sharper tuning in humans. An early study by Harrison et al. (1981) compared Q factors calculated from compound action potential (CAP) recordings in human patients to similar measurements obtained in guinea pig. The CAP is an evoked potential which reflects some properties of ANFs; it is the weighted summed extracellular response of synchronized ANFs at the onset of acoustic stimuli (Goldstein and Kiang, 1958; Kiang et al., 1976; Antoli-Candela and Kiang, 1978). Extrapolating from CAP measurements and from Q factors of single ANFs in guinea pigs, Harrison et al. (1981) concluded that frequency selectivity is higher in humans than in guinea pigs. More recently, based on behavioral (Oxenham and Shera, 2003) and otoacoustic (Shera et al., 2010) measurements, it was concluded that frequency selectivity in humans is higher than previously thought, by a factor of more than two. Finally, cortical recordings in humans also reveal high frequency selectivity (Bitterman et al., 2008). It is controversial whether the high Q values reflect real species differences rather than differences in experimental paradigm (Ruggero and Temchin, 2005; Eustaquio-Martin and Lopez-Poveda, 2011). Comparison of tuning estimates from otoacoustic emissions and single ANF in macaque monkey support the correspondence between the two methods and suggest that frequency tuning in old-world primates may be superior to that observed in more commonly studied species (Joris et al., 2011).

CAP recordings provide an electrophysiological means to compare tuning across species. In order to make predictions of the tuning at the single neuron level in humans, it is necessary to first assess the quantitative relationship between CAP and single unit measures in species where both can be measured or are available. To achieve this goal in the present study, we adopt a forward masking procedure (Oxenham and Shera, 2003) to measure Q factors from CAP recordings in two species for which tuning in ANF is well characterized, but for which little data on CAP tuning is available (Ruggero and Temchin, 2005).

METHODS

Surgical Preparation

The experiments were conducted using 6 adult cats and 9 chinchillas of either sex. All procedures were approved by the KU Leuven Ethics Committee for Animal Experiments and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Animals were briefly examined for clear external and middle ears. Cats were anesthetized with a mixture of ketamine (20 mg/kg) and acepromazine (0.2 mg/kg) administered intramuscularly. A venous cannula was then placed to infuse lactated Ringer's solution and sodium pentobarbital to maintain a deep state of anesthesia. A cannula was inserted into the trachea. Chinchillas were anesthetized with an intramuscular injection of ketamine (20 mg/kg), xylazine (3 mg/kg) and atropine (0.04 mg/kg). Supplementary doses of ketamine or intraperitoneal injection of pentobarbital were administered to maintain a surgical level of anesthesia.

The left pinna was surgically removed and the auditory bulla was exposed and opened. Through the opening in the bulla a silver Teflon insulated wire with ball electrode was inserted and placed at the round window. The electrode lead was glued to the bone and the bulla opening was closed again with ear impression compound (Microsonic). Two silver wire electrodes were placed in the skin: a reference electrode at the nape of the neck and a ground electrode near the contralateral bulla. During the measurements the animals were kept warm with a homeothermic blanket (Harvard). The experiments were conducted in a double-walled soundproof and faradized room (Industrial Acoustics Company).

Stimulus generation

Stimuli were generated with custom software and a digital sound system (Tucker-Davis Technologies, system 2, sample rate: 125 kHz/channel) consisting of a digital-to-analog

converter (PD1), a digitally controlled attenuator (PA5), a headphone driver (HB7) and an electromagnetically shielded acoustic transducer (for chinchilla, Etymotic Research, ER1, 20 Hz - 16 kHz; for cat a dynamic electro-acoustic transducer, Radio Shack, 20 Hz - 50 kHz). The transducer was connected with plastic tubing to a custom earpiece coupler. In cat the coupler was fit in the transversely cut ear canal; in chinchilla the coupler was sealed with ear impression compound to the bony external auditory meatus. The acoustic system was calibrated *in situ* near the eardrum, through the custom coupler, with a calibrated probe-microphone (*Brüel & Kjær,* type 4192, ½-inch condenser microphone and conditioning amplifier Nexus 2690).

Signal sampling

The cochlear responses were differentially measured (low noise differential preamplifier, Signal Recovery, Model 5115) between the signal electrode at the round window and the reference electrode in the nape of the neck. The ground of the differential preamplifier was connected to the electrode at the contralateral mastoid. The signal was filtered (100 Hz - 10 kHz, cut-off slopes 12 dB/oct.) and amplified with another external amplifier (Stanford Research, SR560) to a total gain of x10 000. All of the relevant signals including the filtered cochlear response, the stimulus waveform, and synchronization pulses at the start of every probe were visualized on an oscilloscope (LeCroy, WaveSurfer 24Xs), sampled and digitized with an ADC (TDT, RX8, ~100 kHz/channel, max. SNR 96 dB), and stored on disk for further analysis.

The amplitude of the CAP was measured between P1 (first positive peak) and N1 (first trough) as shown in the inset of Fig. 4 (the remainder of that figure will be discussed later in RESULTS). Whenever P1 was not clearly defined, the amplitude was measured between P2 and N1, and when P2 was not clearly visible, between the positive and negative maxima. Before the amplitude was measured, the CAP-response was de-noised with a wavelet filter. At low probe frequencies (<2 kHz) an additional sharp notch-FIR-filter at twice the probe frequency was used to remove the neural phase-locked response. The measured responses contained significant background noise uncorrelated to the stimulus; at low stimulus intensities the background noise exceeded the desired response signal. To increase the response signal-to-noise ratio (SNR) we used signal averaging, typically 512 or 1024 repetitions, depending on the background noise level.

Stimulus Paradigm

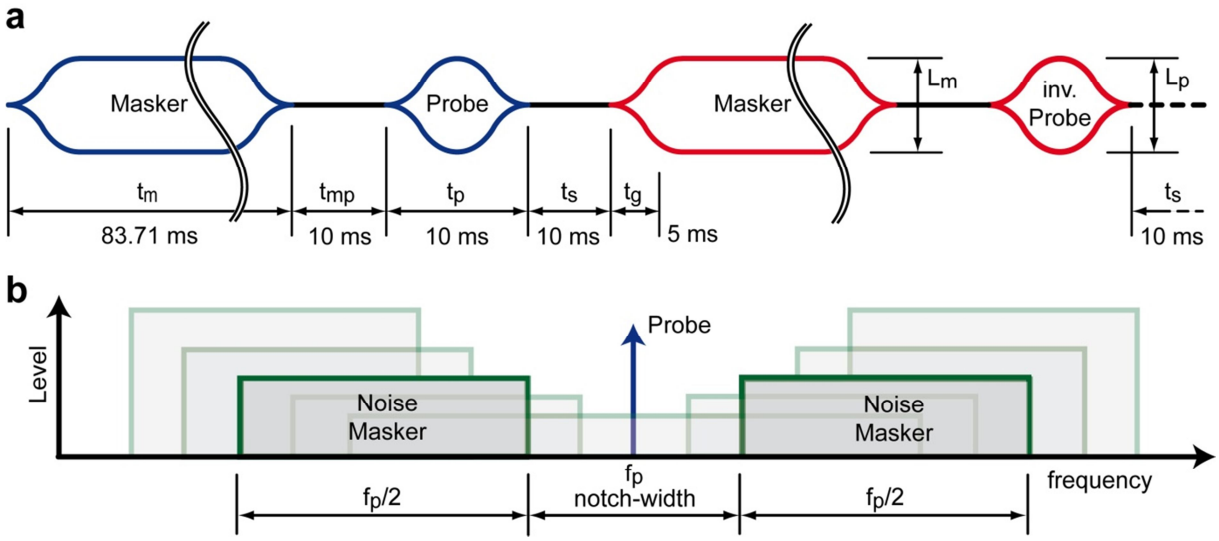


Figure 1. Representation of the stimulus in the time domain (a) and frequency (b) domain with the following parameters: masker duration, t_m; probe duration, t_p; masker-probe interval, t_{mp}; probe-masker interval, t_s; gating time, t_g; masker spectrum level, $L_{m;}$ probe level, L_p; and probe frequency f_p. a: Each complete representation consists of two identical masker-probe sequences with alternating polarity between first and second probe. b: The notched-noise consists of two Gaussian noise bands with fixed bandwidth $f_p/2$. The notch-width and masker level (L_m) are variable.

Schematic representations of the notched-noise forward masking (NNFM) paradigm in the time and frequency domain are shown in Fig. 1. The general principle is as follows. A notched-noise masker of variable level was followed by a 10 ms tonal probe at a fixed level (Fig. 1a). We varied the notch-width of the masker and searched for the masker level required to suppress the CAP response to the fixed probe tone by a given percentage, from hereon called the masking criterion. (A full description of the protocol is given in the next section). To reduce spectral splatter due to the on- and off-switching of the stimuli, probe and masker were gated with 5 ms raised-cosine ramps. The 10 ms probe thus never reached steady state; it smoothly ramped up and down. The polarity of the probe tone alternated in successive presentations. The cochlear microphonic follows the polarity of the stimulus and is largely symmetrical, so that it is almost completely removed by averaging the responses to stimuli with alternating polarity (Dallos, 1973). In contrast, the CAP - the signal we want to measure - is mainly from neural origin and is largely preserved after averaging.

Fig. 1b depicts the idealized frequency spectrum of the stimulus. The notched-noise masker consisted of two Gaussian noise bands generated in software in the complex frequency domain and designed with infinitely steep edges (brick-wall shape). The real and imaginary parts of the spectral components were generated with a Gaussian random generator (*randn*) in MATLAB, spectrally calibrated and converted to the time domain with a discrete

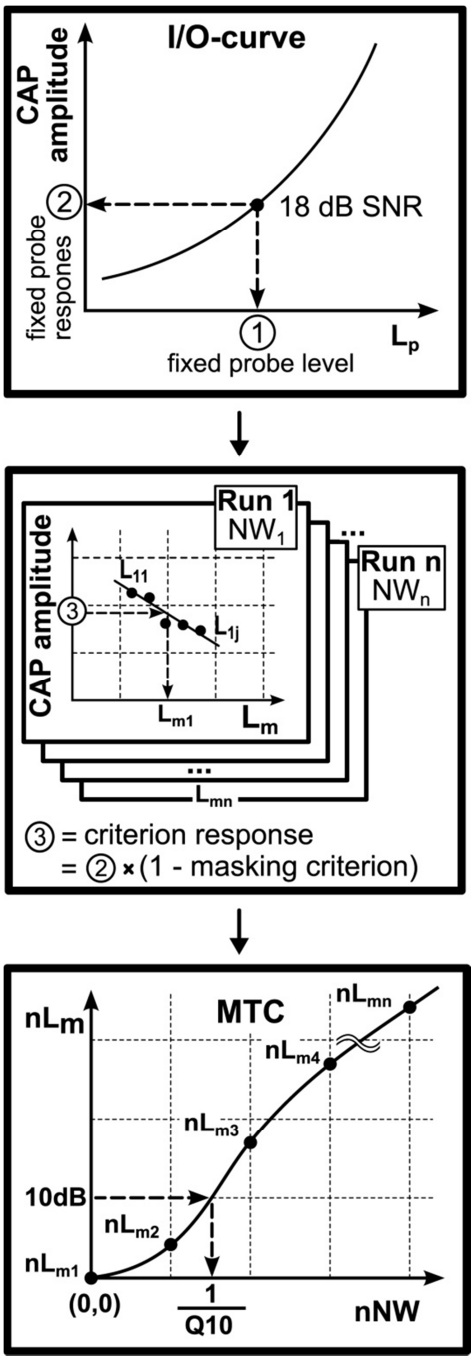
Fast Fourier Transformation (*FFT* function in MATLAB). For convenience, the amplitudes for the masker and probe signals were calculated before signal gating and before calibration. Note that a) the probe levels with gating are 4.3 dB lower than those without gating and b) the masker level is expressed as an overall level (total noise band), and not as a "masker spectrum level" (intensity level within a 1 Hz-band). The conversion factor between the overall level and masker spectrum level is $-10 \cdot \log_{10}(BW)$, where BW = f_p in this study. Spectrally, the noise bands had equal bandwidth and straddled the probe frequency (f_p), so as to create a notch symmetrically spaced on a linear frequency axis. The masker notch-width and level were the main experimental variables. The duration of the masker was set to exceed the settling time of the adaptation in the masker response. In initial pilot experiments we used 300 ms, but after validation we shortened the duration to 83.71 ms to speed up data collection. For a few test cases (not included in the results) we shortened it further to 33.71 ms, which resulted in Q values that were not significantly different compared to longer masker durations. Masker and probe were separated by a silent interval t_{mp}, so as to reduce overlap between the responses from the masker and the probe. Successive presentations of probe-masker combinations were separated with the interval t_s. The values of the intervals t_{mp} and t_s were adjusted in initial experiments, and were fixed at 10 ms for the remaining experiments.

Experimental protocol

Masking tuning curves (MTCs) were constructed with a predetermined masking criterion by varying the notch-width (NW) and masker level (L_m) while keeping the probe constant in level and frequency. Within a run, NW was held constant and L_m was varied in search for the probe response associated with the predetermined masking criterion, called the "criterion response". The L_m values obtained at runs for different NW values resulted in one MTC.

The procedure to extract the MTC consisted of the steps schematically illustrated in 2 (top panel). We obtained an input-output (I/O) curve in quiet i.e., the probe level (L_p) was varied in level and the corresponding CAP amplitudes were measured without the masker. From these responses a suitable probe level was determined. The choice of this level has to meet two conflicting demands. On the one hand it should be low to obtain Q values as close as possible to CAP response threshold for better comparison with Q values from frequency threshold curves of single ANFs. On the other hand, it should be high enough to yield a reasonable SNR without requiring many measurement repetitions. For pragmatic reasons we used the SNR of the CAP response (signal and noise in V_{rms}) as *a figure of merit*, and aimed at 18 dB SNR. To obtain a response with such SNR, a stimulus probe level was required of

approximately 15 dB above the CAP-threshold level, which is defined as the (averaged) probe level that produces a CAP response equal to the background noise (≡ 0 dB SNR). We refer to this probe level as the "fixed probe level" (① in Fig. 2) and to its response as the "fixed probe response" (② in Fig. 2). The fixed probe level is maintained throughout the different runs needed to collect data to construct a MTC. Note that the SNR value was chosen to have enough headroom to cover the smaller amplitudes of the masked responses e.g., a masking criterion of 33% reduces the SNR by approximately 3.5 dB.



Next (middle panel Fig. 2), we searched for the masker levels that suppressed the response-amplitude at the fixed probe level with the masking criterion. In initial experiments the masking criterion was chosen at 50%, i.e., we determined L_m that reduced the response-amplitude to the fixed probe by half. In later experiments the masking criterion was refined to 33% (see RESULTS). In the first run (Run 1) we presented a spectrally continuous masker (no-notch condition) and varied the masker level L_m in small steps to search for the level, e.g., L_{m1}, that resulted in the criterion response as estimated by simple linear regression. In the other subsequent runs the masker levels (L_{m2}, L_{m3}, …, L_{mn}) were determined in the same way with the same probe and criterion conditions, but with increasing masker notch-widths.

In the third and final step (bottom panel Fig. 2), a plot of masker levels for different notch-widths (L_{m1}, L_{m2}, …, L_{mn}) constitutes the MTC, from which

Figure 2. Schematic representation of the procedure for constructing MTCs. From an I/O curve a *fixed probe level* ① and corresponding response ② were chosen. Subsequently, a *criterion response* amplitude ③ is calculated by applying the *masking criterion* to the *fixed probe response* ②. Masked probe responses from a varying masker level (L_{11}, L_{12}, …, L_{1j}) are then obtained for consecutive runs with increasing notch-widths. For every run, simple linear regression is applied on the masked probe responses to determine the masker levels (L_{m1}, L_{m2}, …, L_{mn}) resulting in the *criterion response* ③. Finally, the MTC is estimated with polynomial (linear) regression and Q_{10} is extracted.

the Q factor is obtained. To facilitate comparison across MTCs, we normalize both abscissa and ordinate. Masker levels at different notch widths are normalized by subtracting the masker level for the no-notch condition L_{m1} and the notch-widths are divided by the probe frequency (f_p). Thus, the first data point (NW_1, L_{m1}) with the no-notch condition is positioned at the origin of the normalized MTC. For clarity, the names of the normalized parameters are indicated with the letter 'n' at the beginning of the corresponding label (e.g., $nNW_2 = NW_2 / f_p$ and $nL_{m2} = L_{m2} - L_{m1}$). The MTC-curve is estimated by second order polynomial regression. This is illustrated in Fig. 3 where the data points from all runs, with different notch widths and masker levels, are plotted (open circles). For a better view the axes are rotated relative to Fig. 2 (bottom panel); the extra axis (x-axis) is the probe response. The dashed curves are fits, obtained with polynomial linear regression, through the responses of one run (i.e., one notch width). The intersections of these fits with the criterion response (here 2.6 µV) are indicated with crosses. The MTC is obtained as a fit through these crosses using polynomial linear regression (solid line in Fig. 3).

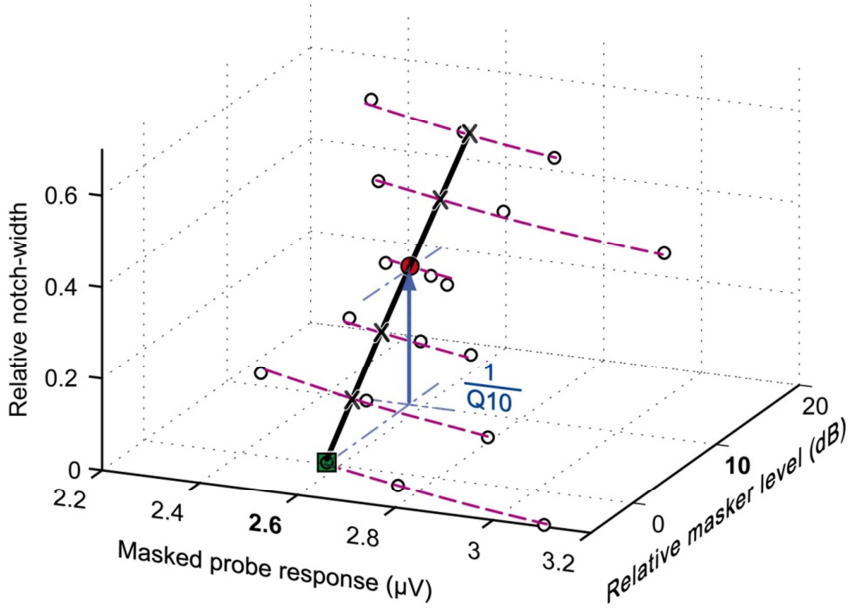


Figure 3. An example of a normalized MTC (thick line) with the masked probe response as extra dimension, here presented on the x-axis. The independent variables are relative notch-width and relative masker level, presented on the y- and z-axis, respectively. The purple dashed curves were obtained by polynomial regression from masked probe responses (open dots); the parameter is the notch-width. The (black) thick curve is the MTC obtained with a similar regression method, but from data points on the (purple) dashed curves at the response criterion (~2.6 µV; black crosses, red circle and green square). The Q_{10} factor is graphically illustrated by the blue arrow (projection of the MTC between red filled dot and green filled square on the axis of the relative notch-width); it is the reciprocal of the normalized notch-width at 10 dB relative masker level.

In this study sharpness of tuning is quantified using two Q factors: the Q_{10} and the Q_{ERB}. Since sharpness of tuning obtained with the notched-noise paradigm is compared with that from single nerve fibers obtained with tones, the Q factors were extracted in two ways. First, directly from the iso-response data, and secondly from a calculated auditory filter (Patterson, 1976), which is the representation of an equivalent filter for tonal maskers.

Direct extraction method: For this method the Q_{10} is defined as f_p/NW_{10dB} where f_p is the probe frequency in Hz and NW_{10dB} is the notch-width at 10 dB above the nadir of the MTC. This notch-width is graphically illustrated by the blue arrow in Fig. 3. The Q_{ERB} is the ratio of the probe frequency to the MTC's equivalent rectangular bandwidth (ERB). The ERB is an approximation of the bandwidth of the filter, modeled as a rectangular band-pass filter passing the same power. The Q_{ERB} is calculated as

$$Q_{ERB} = \frac{f_p}{ERB} = \frac{1}{\int_{-\infty}^{+\infty} \left(\frac{1}{nMTC}\right)^2 dnNW} \qquad (1)$$

where Q_{ERB} is the dimensionless Q factor; f_p is the probe frequency in Hz; ERB is the equivalent rectangular bandwidth in Hz; nNW is the dimensionless notch-width normalized to the probe frequency and nMTC is the normalized MTC in dB. Because the exact MTC is not known and needs to be estimated, we use a numerical integration technique to approximate the definite integral in Eq. 1. The Q_{ERB} is calculated with the trapezium rule for non-uniform intervals. The following equation was used

$$Q_{ERB} \approx \frac{2}{\sum_{i=2}^{N} (nNW_i - nNW_{i-1}) \left(\frac{1}{nL_{mi}{}^2} + \frac{1}{nL_{mi-1}{}^2}\right)} \qquad (2)$$

where nNW_i are the dimensionless normalized discrete notch-widths; nL_{mi} are the normalized discrete masker levels in dB; and N is the number of runs. Due to time constraints only the most relevant part at the region of the nadir was measured and used in the calculation of the Q_{ERB}.

When enough data points are available, the Q_{ERB} is a more robust measure for the sharpness of tuning than the Q_{10}. The Q_{ERB} is calculated from all available data points, whereas the Q_{10} is derived from the data points at 10 dB above the nadir of the MTC. This makes the Q_{ERB} less prone to local variations and more informative regarding the curvature of the slopes. Nevertheless, to limit the total time span needed to obtain a single Q value, we optimized our measurements more towards Q_{10} than to Q_{ERB} estimates. To achieve this, in some cases a more direct method was used where NW_{10dB}, the notch-width at the corresponding masker level 10 dB above the nadir, was sought at the masking criterion; the

Q_{10} is then calculated as the reciprocal of the normalized NW_{10dB}. We refer to this method as the "fast method".

Extraction from an auditory filter: The Q factors were calculated in a similar manner as for the direct extraction method, but from estimated tonal auditory filters commonly used in psychophysical experiments. The auditory filter shapes were obtained with the same fitting procedure as used in a recent psychophysical study (Oxenham and Shera, 2003), but without correction for the middle ear transfer function and off-frequency listening. This fitting procedure is based on the power spectrum model of masking (Patterson, 1976) assuming a rounded exponential (ROEX) filter shape (Patterson et al., 1982; Glasberg et al., 1984). Here, a symmetrical ROEX(*p,w,t*) function (Equation 3 in Oxenham and Shera, 2003) with 3 independent parameters was used, with two equal filter sides. The equation for one of the symmetric sides of this ROEX filter is equal to

$$W(g) = (1-w)(1+pg)\exp(-pg) + w\left(1+\frac{pg}{t}\right)\exp\left(-\frac{pg}{t}\right) \quad (3)$$

where W(g) is the filter power weighting function, g is one half of the discrete normalized notch-width ($nNW_i/2$), w is the weighting factor between the first and second term, p is a measure for the slope of the left term and t is the ratio between the slopes of the left and right terms. From the power spectrum model a fourth, dependent, parameter K (dB) was extracted, which represents the masking "efficiency" corresponding to the predefined masking criteria, and is defined as the ratio of the powers of the probe and the amount of masker (spectral noise) passing the filter (W) necessary to comply with the masking criterion. For clarity, we indicate the sharpness of tuning obtained from these filters with the prefix ROEX.

In a few animals we also studied the asymmetry of MTCs obtained with a single noise band positioned above or below the probe frequency. In those instances, the two MTCs obtained with single noise bands were merged and represented as a V-shaped MTC; we refer to these as single-sided MTCs (ss-MTC).

RESULTS

In this study we focus on the sharpness of tuning of CAP responses in cat and chinchilla. We calculate Q_{10} and Q_{ERB} factors from iso-response tuning curves derived from masked CAP responses recorded at the round window using a NNFM-paradigm. The most complete datasets were obtained in cat. The probe frequencies varied from 0.5 to 14 kHz for cat and from 1 to 12 kHz for chinchilla. The sequence of probe frequencies tested was

arbitrary. In the course of an experiment, repeated measurements were occasionally made as a control: in those cases average Q factors are calculated and reported. We strove to use probe levels as low as practical. In some measurements the probe level was systematically increased to explore the dependence of Q_{10} on level.

Fixed probe level

The first step was to determine a fixed probe level for the remainder of the procedure. Examples of typical CAP responses from cat to unmasked tone pips are depicted in Fig. 4, for a probe frequency of 8 kHz.

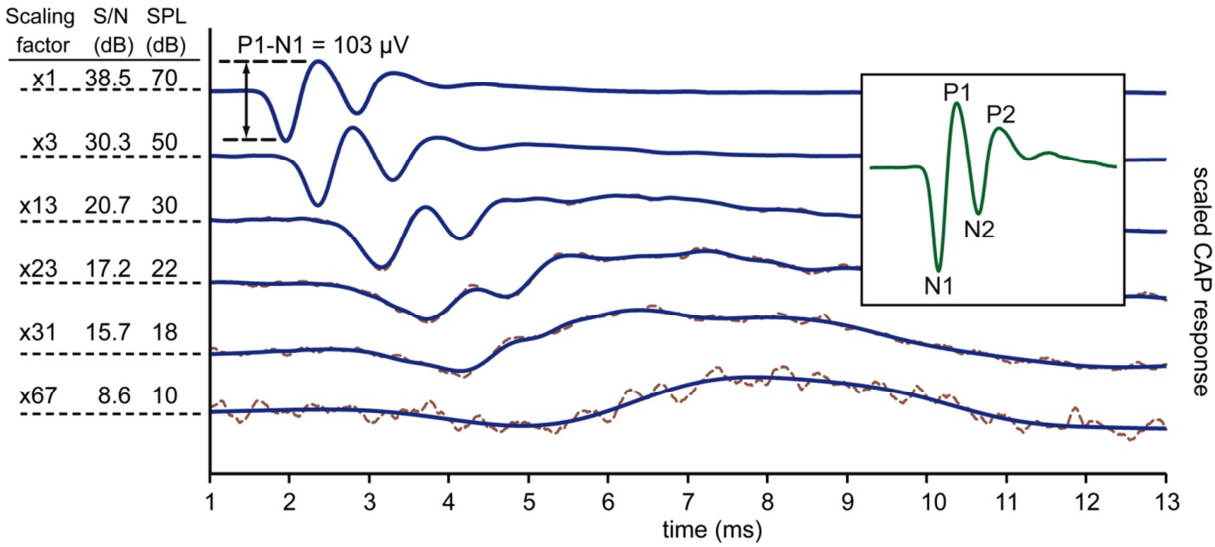


Figure 4. CAP-responses extracted from the alternated round window-responses for different probe levels, dashed signals. The solid blue traces are the waveforms de-noised with a custom wavelet filter; the curves are offset and the amplitudes scaled. Table at the left indicates the scaling factors, the SNR (response and background in V_{rms}) and the probe levels. The probe frequency is 8 kHz and response amplitude for 70 dB is 103 μV_{pp} (peak to peak). Inset: the CAP response at 70 dB, with the different positive and negative peaks. The time-axis is not compensated for the system delay, e.g. acoustical transmission tubes (0.3 ms).

Traces for 6 different probe levels, increasing from bottom to top, are shown. The amplitudes are scaled: the table to the left gives the scaling factor, SNR and probe level for the corresponding responses. In this example, 22 dB SPL was chosen as the fixed probe level, based on the corresponding SNR of 17.2 dB, which is close to our predefined target value of 18 dB. Note that at high levels, P1 and N1 (as well as other components such as P2, N2) are clearly defined so that the CAP response can be simply defined as their amplitude difference. However, at low levels, a broad maximum rather than a distinct P1 component was often seen (lower traces in Fig. 4) and was used to quantify the response amplitude. In these cases it is

impossible to know whether this absolute maximum corresponds to P1, P2 or the summating potential or a mixture thereof. In practice, the exact choice of response metric is not critical because our paradigm uses iso-response conditions and the intrinsic shape of the CAP response does not change much with notch-width and masker level.

Masking criterion

Earlier studies (Eggermont, 1977; Harrison et al., 1981) used a variety of masking criteria. In our initial experiments we used a masking criterion of 50%, but subsequently settled on 33% which we found to be optimal, as illustrated in Fig. 5.

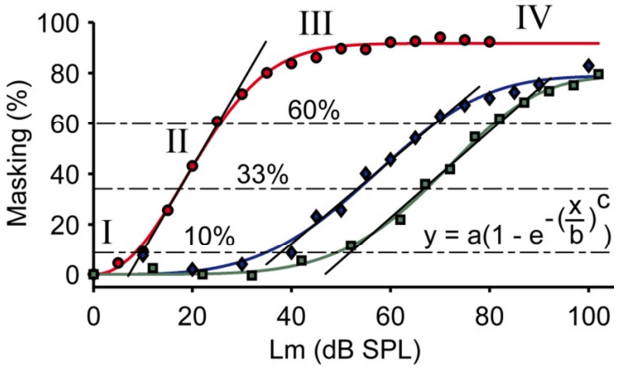


Figure 5. Masking intensity curves, amount of masking as a function of masker level, for different conditions in 2 different cats. The masking curves are fits of cumulative Weibull distribution functions (formula included in figure). The conditions for the curves are, from left to right: red filled dots, L_p = 40 dB SPL, f_p = 10 kHz, nNW = 0, a=91.68, b=24.89, c=2.129, animal 1; blue diamonds, L_p = 67 dB SPL, f_p = 2 kHz, nNW = 0.44, a=78.89, b=62.62, c=3.668, animal 2; and green squares, L_p = 60 dB SPL, f_p = 6 kHz, nNW = 0.34, a=78.72, b=75,29, c=5.031, animal 2. Four different regions can be identified: I, region with almost no masking; II, region with a steep linear growth (on semi-log plot, solid lines), i.e., between ~10 – 60 % masking (upper and lower horizontal dashed-dotted lines); III, transition region with a much smaller growth; and IV a region clipped to maximal masking. The middle of region II is the best region for the masking criterion, chosen at ~33% (middle horizontal dashed-dotted line).

This figure shows 3 masking intensity curves in cat for three different conditions (see caption Fig. 5). The normalized masking is plotted as a function of masker level. The curves show 4 distinct regions: Region I: no masking; Region II: steep linear growth of masking between ~10 – 60 % (upper and lower horizontal dashed dotted lines) on a semi-log scale; Region III: transition region, with a slope significantly smaller than in region II; Region IV: saturation of masking, the region is only visible for the leftmost curve (red filled dots). Region II shows a linear relationship between level and masking in the three curves, which differ in frequency, notch-width and probe level. We therefore chose 33% as the optimal level for our masking criterion (middle dashed line) and used simple linear regression. There is a

clear difference in slope between the two curves with higher probe levels (blue diamonds and green squares, slopes of 1.7 and 1.8 dB^{-1}, respectively) and the one with low probe level (red dots, 3.4 dB^{-1}). The masking intensity curves are similar to those for single fiber measurements (Harris and Dallos, 1979); the red curve in our example with low probe level is almost identical with the same ascending slope to the normalized curve in their Fig. 9.

Masking recovery

As a final check on the chosen parameter values in our NNFM paradigm i.e., t_m, t_{mp} and t_s, we verified the (fast) recovery from masking. Fig. 6 illustrates the recovery of a 10 kHz, 10 ms, probe tone from a preceding no-notch noise masker with a duration of 83.71 ms (t_m); t_s was chosen to allow full (fast) recovery. The recovery function is measured in successive measurements where t_{mp} is increased from 0 to 30 ms in steps of approximately 5 ms while the masker to masker interval was held constant. The data points are well described by a recovery curve based on an exponential fast recovery model used for forward masking in single auditory nerve fibers (Harris and Dallos, 1979). The exponential fast recovery model predicts that >99% of the probe response would recover within the next probe to probe interval of 113.71 ms ($t_p + t_s + t_m + t_{mp}$), which validates the current chosen parameter values.

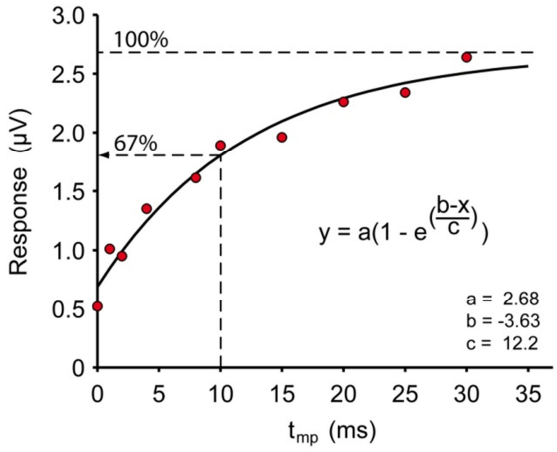


Figure 6. An example of a recovery curve derived with the NNFM paradigm. The amplitudes of the masked responses (red dots) are plotted as a function of ten different masker-probe intervals (t_{mp}) ranging from 0 to 30 ms. The conditions of the measurement were: nNW = 0, $L_m = L_p = 40$ dB SPL, $f_p = 10$ kHz. The solid curve is a fit through the data points using the formula in the figure with y the response and x the masker-probe interval. The parameters a, b and c were determined using the Gauss-Newton method for non-linear regression. The upper dashed line is the unmasked amplitude limit. The lower dashed line shows the fitted response value at $t_{mp} = 10$ ms, which is the value used in our standard paradigm.

MTCs, Q_{10} and Q_{ERB}

Fig. 7 presents 3 sets of MTCs obtained in 3 different cats for various probe frequencies. The MTCs are reconstructed by straight lines between the data points. All MTCs show a monotonic increase with notch-width and can have a slightly concave or convex curvature. In the 3 animals, there is a consistent increase in slope with increasing probe frequency which is a good indication that the CAP responses are generated by single units with CF close to the probe frequency.

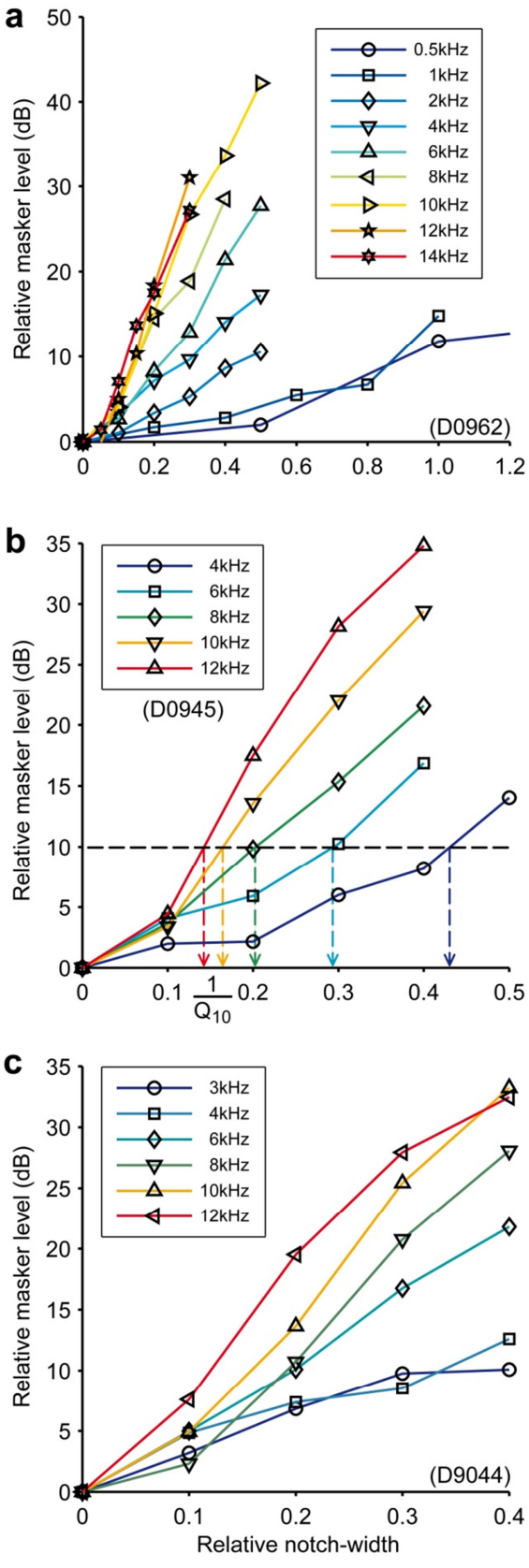


Figure 7. Normalized MTCs for 3 different cats (a, b and c). The relative masker level is plotted as a function of relative notch-width for a range of probe frequencies. The extraction of the estimated Q_{10} from the MTCs is illustrated in **b**; it is the reciprocal of the intersection of the dashed line at 10 dB relative masker level with the MTC. The masker reference levels (overall level in dB) for the MTC's are presented in Table 1, indicated by "Ref L_m".

The notched-noise paradigm yields one-sided masking tuning curves and, in its standard configuration with symmetrical notch, does not allow separate examination of the part of the filter above and below the probe frequency. One of the assumptions made, validated below in *Single sided noise maskers*, is that the MTC, at low probe level, is symmetric around the nadir i.e., both noise bands produce the same amount of masking for a given notch width. As illustrated in Fig. 7b, the intersection of these one-sided tuning curves with a horizontal line at a relative masker level of 10 dB provides an estimate of the Q_{10} for notched-noise maskers.

Fig. 8 shows the two indicators of sharpness, Q_{10} (Fig. 8a) and Q_{ERB} (Fig. 8b), plotted on a linear frequency axis for different cats (indicated with different symbols, see legend). In both figures the filled dots, connected with dashed lines, are the average values across animals. The mean trend line of the Q_{10} (Fig. 8a, dashed line) shows a fairly linear relation with probe frequency. For some probe frequencies, measurements are available for only a single animal and are thus statistically less reliable. For example at 14 kHz the mean trend line suggests a decreasing Q_{10} factor beyond 12 kHz, but this is not supported by the within-animal trend

(experiment A0962, ... $< Q_{10,\ 10kHz} < Q_{10,\ 12kHz} < Q_{10,\ 14kHz}$). A trend-predictor more resistant to outliers, missing and scattered data points is the robust LOESS trend line (MATLAB routine *smooth*, method *RLOESS*, span 0.7), indicated in red in both plots. The trend lines show a, monotonically increasing, smoother course than the mean trend lines. The Q_{ERB} trend line (Fig. 8b) shows a clear bend at 4 kHz with a much steeper slope at the lower frequency side.

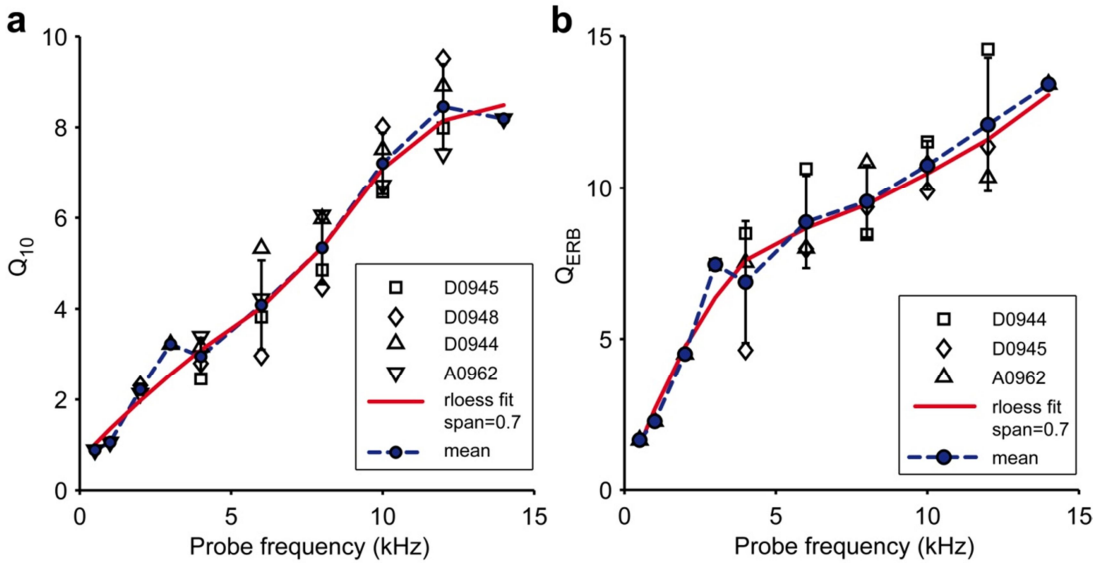


Figure 8. Dependence of Q factors on probe frequency in cats. a: Q_{10} values for 4 cats are plotted against probe frequencies ranging from 0.5 to 14 kHz. The blue filled dots connected with dashed lines are the mean values. The error bars indicate the standard errors of the mean; they are omitted for single data points. The red line is a trend line calculated with a *RLOESS* function (MATLAB). b: same as a, but for Q_{ERB} for 3 cats.

In this study Q factors are obtained with notched-noise maskers whereas Q factors from single fibers are obtained with tones. To make a more appropriate comparison between the results from different maskers/exciters i.e., noise maskers versus tones, the MTCs from 3 cats were converted to auditory filters as commonly used in psychophysical experiments (see *Extraction from an auditory filter* in METHODS). The filter parameters and Q factors obtained from the ROEX filters are given in Table 1. It is clear from this table that for most of the obtained ROEX filter functions the second term in the ROEX(*p,w,t*) is negligible ($w < -100$ dB). Note that the masking "efficiency", K, in the table depends on the specific probe level and masking criterion and that direct comparison between K factors with different conditions is meaningless.

Table 1

Parameters, filter parameters and results derived from the same data as in Fig. 7

Cat 1: D0944					*masking criterion: 50%*			
Frequency	L_p	*Ref* L_m	*p*	*t*	*10 log(w)*	*K*	Q_{10}	Q_{ERB}
3000	36	43	17.1	308.9	-335.7	-1.9	2.2	4.3
4000	28	39	18.4	578.3	-342.7	-5.4	2.4	4.6
6000	12	21	32.7	149.4	-439.2	-0.2	4.2	8.2
8000	13	21	42.8	559.7	-402.9	4.3	5.5	10.7
10000	15	23	48.9	239.9	-447.5	4.3	6.3	12.2
12000	20	25	60.0	5.6	-34.6	7.4	7.7	15.0
Cat 2: D0945					*masking criterion: 50%*			
Frequency	L_p	*Ref* L_m	*p*	*t*	*10 log(w)*	*K*	Q_{10}	Q_{ERB}
4000	50	51	16.8	394.0	-396.2	6.0	2.2	4.2
6000	41	35	24.7	361.7	-365.8	14.4	3.2	6.2
8000	34	34	32.5	124.2	-307.4	8.9	4.2	8.1
10000	29	29	44.1	253.8	-390.6	11.3	5.6	11.0
12000	33	31	57.0	713.4	-49.9	15.5	7.3	14.2
Cat 3: D0962					*masking criterion: 33%*			
Frequency	L_p	*Ref* L_m	*p*	*t*	*10 log(w)*	*K*	Q_{10}	Q_{ERB}
500	48	38	6.5	24.9	-355.4	14.4	0.8	1.6
1000	47	37	8.4	336.5	-251.4	14.4	1.1	2.1
2000	35	41	16.0	461.1	-383.2	6.9	2.1	4.0
4000	35	40	22.5	419.1	-390.9	6.8	2.9	5.6
6000	30	33	33.8	227.1	-193.3	10.9	4.4	8.5
8000	22	27	42.3	158.1	-317.1	7.5	5.4	10.6
10000	20	20	54.2	2.1	-25.3	14.4	6.9	13.5
12000	20	23	63.3	295.3	-357.4	13.6	8.1	15.8
14000	27	28	59.8	5.7	-29.7	13.1	7.6	14.9

The Q_{10} and Q_{ERB} factors from Table 1 are depicted as a function of frequency in Fig. 9. The ROEX-Q_{10} trend line in Fig. 9a shows the same course as the directly obtained Q_{10} trend line (replotted from Fig. 8a), but with slightly lower values. The relationship (ratio) between the ROEX-Q_{10} and the directly obtained Q_{10} trend lines is constant over the considered frequency range with a mean value of 0.89 (STD: 0.09), and is depicted as the black line in Fig. 9a. Since the ratio is small and because of the additional uncertainties in the ROEX-Q_{10} arising from the assumptions (linearity and ROEX filter shape) in the derivation of the auditory filter, we consider the difference between the two curves too small to warrant use of the ROEX filter method. Therefore, in the remainder of this study we only consider the Q_{10} values directly obtained from the MTCs.

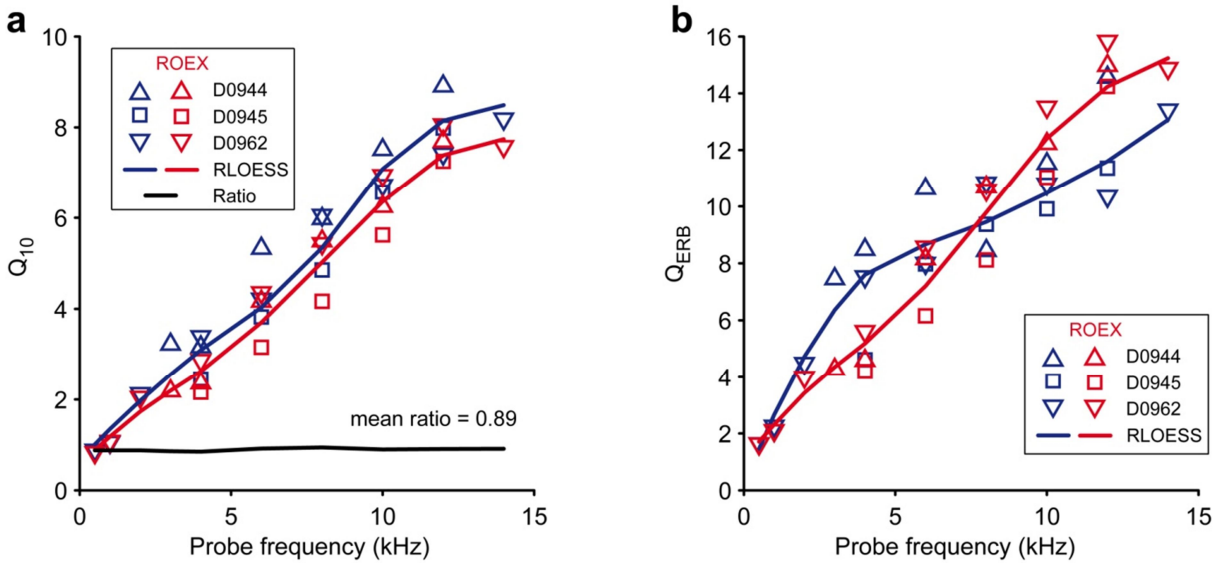


Figure 9. Comparison between directly obtained Q factors (Fig. 8) and Q factors from an estimated ROEX filter function using a power spectrum model for masking. a: Q_{10} and b: Q_{ERB}. Q values for 3 cats are plotted against probe frequencies ranging from 0.5 to 14 kHz. The blue symbols indicate Q values from CAP; the red symbols are extracted from the fitted ROEX filter function. The lines are trend lines calculated with a *RLOESS* function available in MATLAB. The black line in panel a is the relationship between the sharpness of the two trend lines (ROEX-Q_{10} / direct-Q_{10}).

In panel B the values for the ROEX-Q_{ERB} and for the directly obtained Q_{ERB} are shown for the same animals as in panel A. The trend line for the ROEX-Q_{ERB} is besides a proportionality factor identical to the trend line of the ROEX-Q_{10} in panel A. The proportionality factor is equal to 1.95 and is biased by the dominant term in the ROEX(p,w,t) function at the tip of the symmetric auditory filter i.e., $W(g) = (1+p|g|)\cdot exp(-p|g|)$. This value is almost equal to the theoretical value of ~1.945, which is calculated from the following derived equation

$$Q_{RATIO}(\alpha) = \frac{Q_{ERB}}{Q_{\alpha}(\alpha)} = -\frac{\Omega_{-1}\left(-e^{-1}10^{\left(-\frac{\alpha}{10}\right)}\right)+1}{2} \qquad (4)$$

where α means the relative level (dB) associated to the type of the Q factor e.g., α = 10 for Q_{10}; and where Ω_{-1} represents the omega function, which is the inverse relation of the function $f(w) = we^{w}$ for the lower branch. The course of the directly obtained Q_{ERB} deviates significantly from the ROEX-Q_{ERB}; it is further discussed below in the part "*Comparison with single auditory nerve fibers*".

Fig. 10 presents Q_{10} factors as a function of probe frequency for 9 chinchillas. The mean values have a similar increasing trend with probe frequency as in cat, but display much more variability. Despite the large variability, the trend (*RLOESS*) is very similar to the trend for cats. The trends for cat and chinchilla are further discussed below.

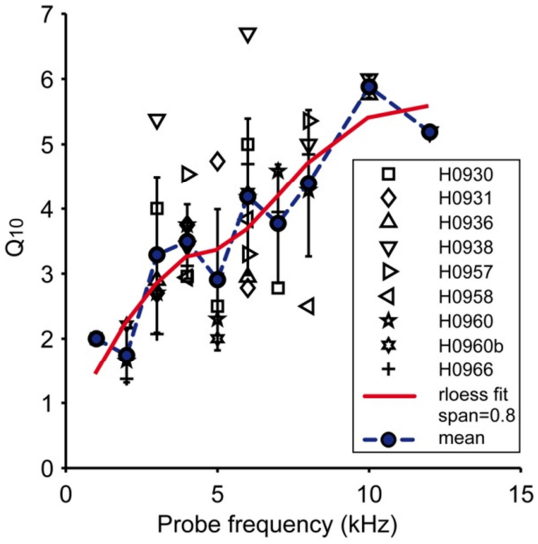


Figure 10. Q_{10} as a function of frequency for 9 chinchillas, the probe frequency ranges from 1 to 12 kHz. The (blue) filled dots connected with dashed lines are the calculated mean values. The error bars indicate the standard errors of the mean; they are omitted for single measured data points. The red line is calculated with a *RLOESS* function (MATLAB).

Dependence on parameter choices

Most studies examining frequency tuning are conducted at threshold or at low probe levels, at which frequency selectivity is highest (Robles and Ruggero, 2001). Due to the nature of the recorded signal and paradigm, compromises had to be made regarding SNR and number of repetitions. Probe levels were never at threshold, but at least 15 dB above CAP-threshold and varied across the measured MTCs. So far, in the comparison between MTCs, we made abstraction of the probe level by normalizing the results. In this section we examine the implication of some of our experimental parameter choices.

Probe Level. We examined the dependence of Q_{10} on probe level. Cochlear frequency selectivity decreases with level: this is well documented for iso-input measurements although it may increase for iso-response measurements (Eustaquio-Martin and Lopez-Poveda, 2011). Fig. 11 presents Q_{10} values as a function of probe level for probe frequencies ranging from 3 to 10 kHz, for 3 different cats.

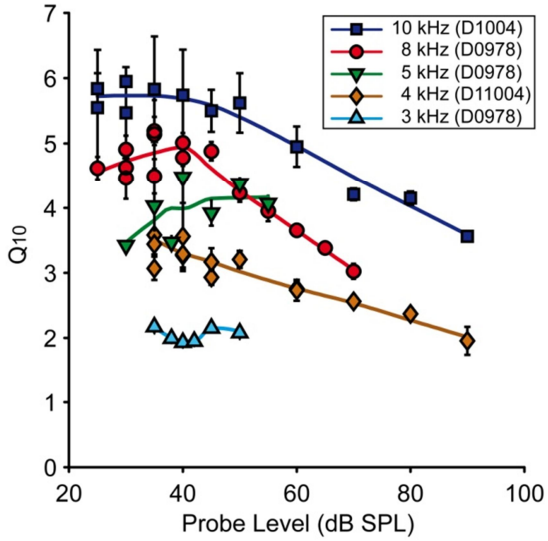


Figure 11. Dependence of Q_{10} factors on probe level for different probe frequencies in cat. Repeated measurements are shown separately; the error bars indicate the standard error of the estimate for regression (SEE). The lines are smoothed trend lines through the scattered data (the trend is obtained with a *RLOWESS* function and smoothed with a cubic spline interpolation (MATLAB). All results were obtained with the fast method.

The trend lines for the most comprehensive data show a low-pass behavior with a cut-off level between approximately 40 and 50 dB SPL. Below the cut-off level the Q_{10} is roughly independent of the probe level. Some trends i.e., for 5 and 8 kHz might suggest a local increasing or decreasing trend with frequency, but these are not significant, given the large standard error of the estimate for regression Above the cut-off level, the Q_{10} declines monotonically with slopes of approximately -0.050, -0.067 and -0.046 dB/dB for 4, 8 and 10 kHz (curves with data points above 55 dB SPL). The probe levels that would typically be used (with the predefined SNR criterion of 18 dB) would be 45, 40, 35, 35 and 40 dB SPL for 3, 4, 5, 8 and 10 kHz, respectively. Thus, the Q_{10} values obtained with our 18 dB SNR criterion represent the highest values that can be measured with this paradigm at all probe frequencies.

We interpret the decline of Q_{10} beyond the cut-off level (Fig. 11) to reflect a change in operation mode (linear to compressive) of the probe. Mechanical measurements of basilar membrane motion to tones (Robles and Ruggero, 2001) reveal an input-output (I/O) relationship consisting of 3 regimes (Fig. 13b): linear (1 dB/dB) regimes at the lowest and highest SPLs, connected by a compressive portion with shallower slope (taken at 0.2 dB/dB).

At low probe levels, both probe and masker are in the linear regime and Q values are independent of SPL. With our masking criterion of 33%, at higher probe levels the probe can be in the compressive regime of the I/O curve while the masker is in the low-SPL linear regime. In this situation, a change (reduction) in response due to a small increase in masker level is greater than a change (increase) in response to the same small increment in probe level. The decline in Q_{10} associated with the compression of the probe is not the result of the corresponding additional increase in probe level, but is related to the associated deformation (widening) of the probe's excitation pattern. As a global effect, to keep the amount of masking constant the notch-width of the masker has to be widened, so the Q value decreases.

growth-of-maskability. The interplay between probe level and masker level is traditionally measured with growth-of-masking (GOM) functions, which describe the dependence of masked threshold on masker level. We examined how the masker level required to generate a criterion response depends on probe level. Following the psychophysical literature (Nelson et al. 2001), we call the resulting functions (Oxenham and Plack, 1997) growth-of-maskability (GMB) functions. According to the reasoning in the previous paragraph, we expect GMB functions to have a slope < 1 dB/dB if the probe level is in a compressive regime while the masker level is in a linear regime. GMB functions for several probe frequencies are plotted in Fig. 12 using a no-notch masker (masker reference level i.e., L_{m1} in Fig. 2). Masker (reference) levels for probe levels above roughly 40 dB SPL, associated to the cut-off level in Fig. 11, show little dependence on probe level. The slopes are smaller than 1 dB/dB (cf. dashed line of equality) and closer to ~0.2 dB/dB (lower dashed line). These small slopes are consistent with our interpretation that the masker is operating in the linear region while the probe is operating in the compressive region, causing the decline in Q_{10} with probe level (Fig. 11). At the lowest probe levels the data points deviate significantly from the course of the other data points; the reason is unclear and is further briefly discussed below in the topic about the GMB model.

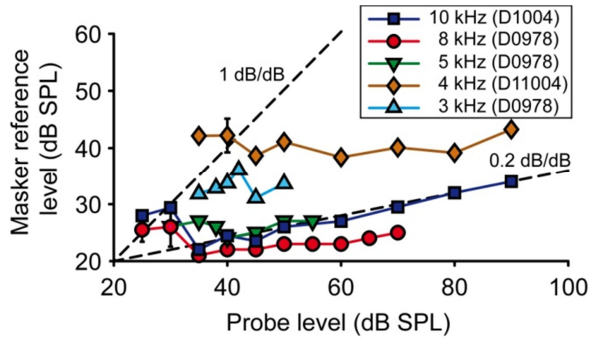


Figure 12. Growth-of-maskability (GMB) functions for the same data as in Fig. 11. The data points are mean values measured at the masking criterion (33%) and masker reference level conditions (no-notch, continuous noise spectrum). The dashed lines indicate a compression exponent of 0.2 dB/dB and 1 dB/dB for equality. The error bars indicate standard error of the mean, when applicable.

For comparison with psychophysical data, it is useful to illustrate our reasoning with a very simple model. Fig. 13a shows GMB curves for different masking conditions. For example, let us assume that for a given masking condition the effective "output" (e.g., the vibration pattern of the basilar membrane) in response to the probe stimulus has to be 15 dB *higher* than the response to the masker. At low stimulus intensities, in the linear regime, this requires a masker level 15 dB below the probe level. At high stimulus intensities, with the probe in the compressive regime, the masker level required is much lower than -15 dB relative to the probe level (e.g., for the I/O function in Fig. 13b this would be a difference of -47 dB for a probe level of 80 dB SPL). The -15 dB blue line in Fig. 13a shows all combinations of probe and masker levels that will generate at 15 dB difference between probe and masker output. If, on the other hand, the masking conditions would be such that the response to the probe stimulus has to be 15 dB *lower* than the response to the masker, the pattern is inversed so that the predicted GMB curves are mirror symmetric relative to the diagonal of equality (red vs. blue curves in Fig. 13).

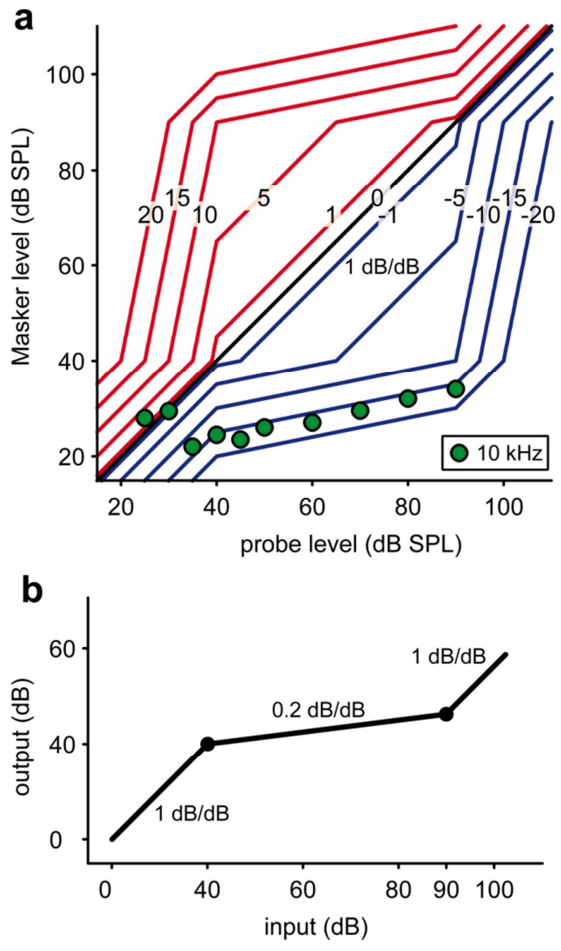


Figure 13. Illustration of the influence of cochlear compression on the GMB. **a**: Simulated GMB curves for different masking conditions. The simplified model assumes that: 1) the probe and masker input levels are subject to the same I/O curve (panel **b)**; 2) masking is a linear operation; and 3) the masking criteria are proportional to the ratios of the excitation patterns (power) from masker and probe. These ratios, in dB, are indicated on the curves in panel **a**. As comparison, the data points of the GMB in Fig. 12 for a probe frequency of 10 kHz are included; the resemblance in absolute values is only a coincidence. **b**: The BM I/O curve (log-log scale) used in the model. The compression coefficients (slopes) and cut-off points (40 and 90 dB) are chosen as suggested values.

The masking conditions (paradigm parameters such as masking criterion, masker-probe interval, masker length, etc.) determine the relative levels required between masker and probe to obtain masking and therefore the expected trajectory of the GMB curves. The black line in Fig. 13a (diagonal of equality) represents the relation when the levels of probe and masker are equal. This situation can be obtained with tonal on-frequency maskers; for an example see listener CP for 6 kHz in Fig. 2 (left) in Oxenham and Plack (1997). For all other curves, the compressive nature of the cochlea introduces regions with slopes deviating from unity. For an idealized I/O-curve, the values of these slopes correspond to the compression exponent (e.g., 0.2 dB/dB or 5 dB/dB). GMB or GOM curves obtained from studies using two-interval forced choice (2IFC) adaptive procedures or using "linear" off-frequency maskers to estimate compression exponents (Relkin and Turner, 1988; Oxenham and Plack, 1997; Plack and Oxenham, 1998) typically have higher masker levels than probe levels and look similar to the red curves in Fig. 13a. In the present paradigm, the masker level needed to achieve the masking criterion of 33% is significantly lower than the probe level and therefore the shape of the curves is reciprocal (blue curves) to GMBs reported in the literature.

The slope of the GMB above the cut-off level (~40 dB SPL) for the 10 kHz data (green dots in Fig. 13a) is well predicted by the model. Below the cut-off level some of the masker levels deviate from the prediction of the model i.e., the lowest data points are off-set and do not show a clear predicted slope equal to 1dB/dB. This observation is also valid for most of the other GMB curves in Fig. 12 having different probe frequencies. Note that these data points were measured at unusually low probe levels, below our quality criterion.

Single sided noise maskers. All previous results were obtained with notched-noise maskers symmetrically positioned with respect to the probe frequency. The assumption here is that, at least at low probe levels, the slopes of the MTCs are symmetrical at both sides. To explore the validity of this assumption we compared MTCs obtained using stimuli with single- and (symmetrical) double-sided noise maskers, hereafter named ss-MTCs and ds-MTCs, respectively. Fig. 14 depicts 4 MTCs, with ss-MTCs and ds-MTCs at frequencies of 4, 5 and 8 kHz, obtained from 2 different cats. The abscissa shows the normalized spectral distance between the closest edge of the masker and the probe. The ss-MTCs are obtained by joining two independently measured MTCs obtained with a left and a right single-sided noise masker: negative relative masker-probe distances indicate a masker band below the probe frequency ($f_m < f_P$) while positive values indicate a masker band above the probe frequency ($f_m > f_P$). For the standard double-sided masker, we plot the MTC twice, mirror symmetrically. Note that in this case the spectral distance equals half of the notch-width. For visualization, left and right sides were fit by a trend line and vertically aligned at the origin.

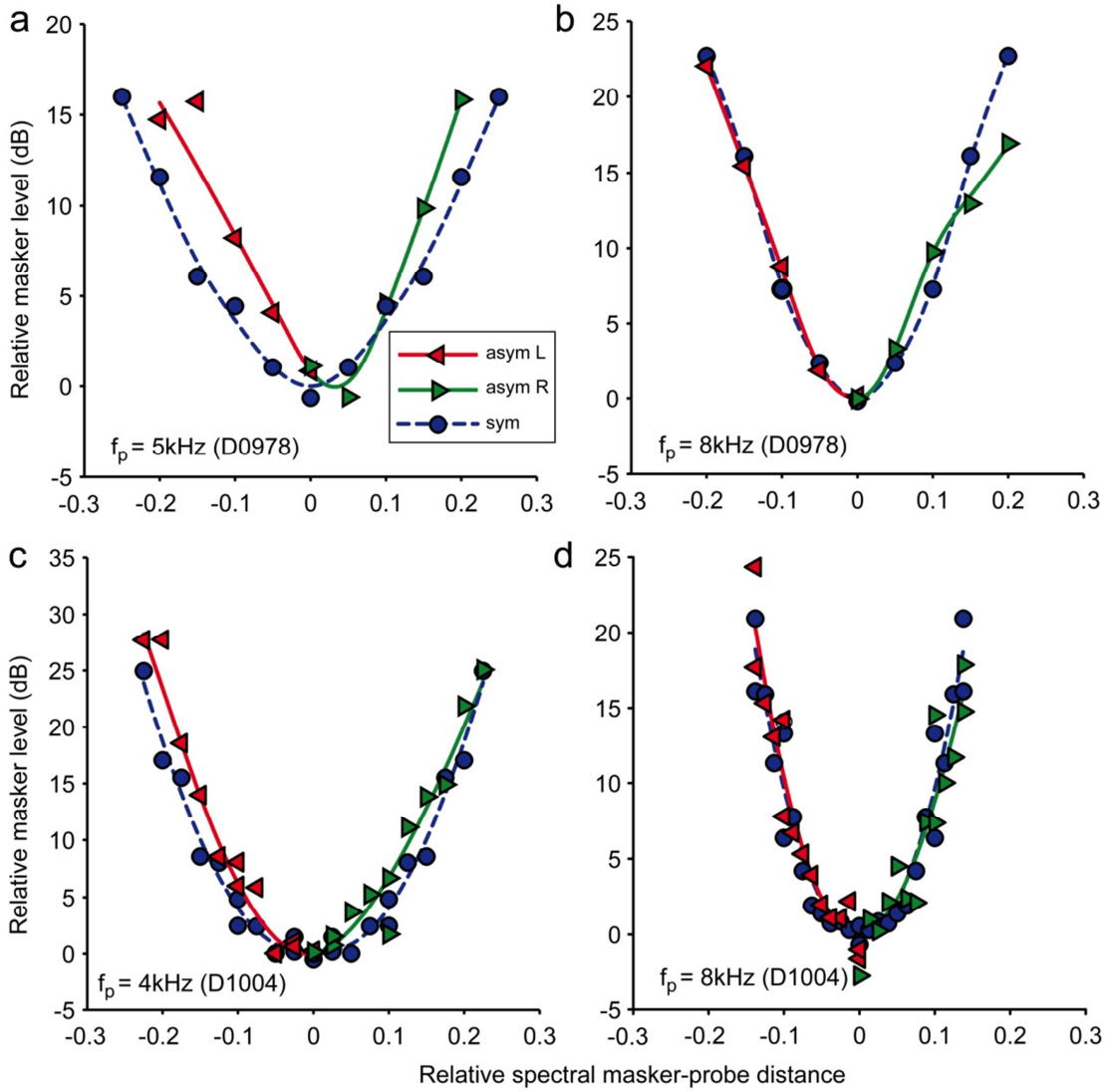


Figure 14. Comparison between left/right single-sided and (symmetrical) double-sided noise masker MTCs. The measured probe frequencies are: **a**: 5 kHz (L_p=35 dB), **b**: 8 kHz (L_p=30 dB), **c**: 4 kHz (L_p=50 dB), **d**: 8 kHz (L_p=40 dB). The trend lines are obtained with *RLOESS* and spline smoothing (MATLAB); the dashed blue line and blue circles denote the standard double-sided condition, the red line and left triangles denote the left single-sided masker condition and the green line and right triangles denote the right single-sided masker condition. The masker reference levels are: **a**: asym L=31 dB, asym R=32 dB, sym =28 dB; **b**: asym L =26 dB, asym R=28 dB, sym =23 dB; **c**: asym L =30 dB, asym R=37 dB, sym =30 dB; **d**: asym L =26 dB, asym R=40 dB, sym =31 dB.

To minimize drift and variability between responses to single- and double-sided noise maskers, all measurements with the same notch-widths were interleaved and recorded in one continuous measurement. The ds- and ss-MTCs do not show large differences. Particularly for single-sided maskers above the probe frequency, the trend line is very similar to that of the

double-sided maskers. For single-sided maskers below the probe frequency, the data points are slightly elevated at the lowest probe frequencies (red curves in Fig. 14c and a, 4 and 5 kHz), but the slopes of the trend lines are similar.

Comparison with single auditory nerve fibers

The main goal of our study was to develop an electrophysiological method to estimate sharpness of tuning from mass cochlear potentials. In this section, we compare our estimates with those for single ANFs. The dots in Fig. 15 show neural Q_{10} values for ANFs available from previous studies in our laboratory. As observed in many previous studies, not only in ANFs but also behaviorally and in the CNS (Oxenham and Shera, 2003; Shera and Guinan, 2003; Sayles and Winter, 2010), Q_{10} increases approximately linearly with CF on a log-log scale. Interestingly, again in agreement with many previous ANF studies in different species (guinea pig: Evans, 1972; Sayles and Winter, 2010; cat: Liberman, 1978; squirrel monkey: Rhode et al., 2010; macaque monkey: Joris et al., 2011), the distribution levels off or even bends over at its upper end i.e., there is no further increase in Q_{10} values at the highest CFs. To our knowledge, the mechanism(s) behind this saturating relationship have not been addressed.

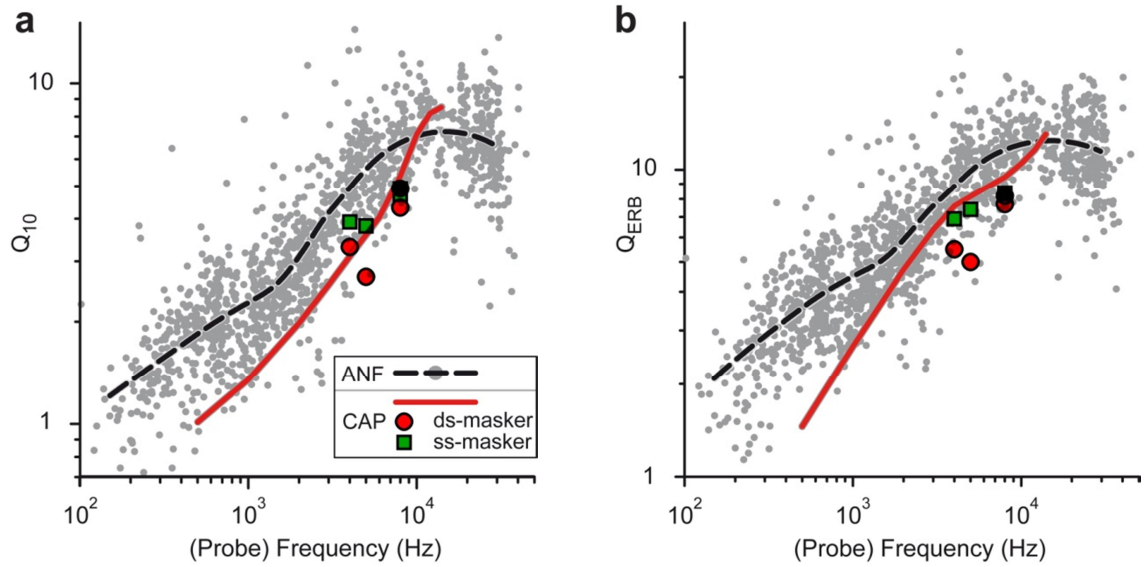


Figure 15. Comparison between Q factors from CAP (red lines, same data as in Fig. 8 but on a log-log scale) and single ANFs (gray dots, 1200 units, 28 cats, CF: 0.1…<50 kHz; threshold level: range from -10 ... 67 dB, spontaneous rate 0 ... 136 s^{-1}, unpublished data) as a function of frequency for cat. a: Q_{10} and b: Q_{ERB}. For ease of comparison, smoothed AN trend lines (dashed lines, *RLOWESS*, MATLAB) are included in both panels. Additional Q_{10} factors from the accurate MTCs in Fig. 14 with double-sided and single-sided maskers are included as the orange circles and green squares.

Fig. 15a compares the Q_{10} CAP trend line (red, same as in Fig. 8a) with the trend of the ANFs (dashed black line). Note that the CAP trend line is located within the cluster of ANF data points. For low and middle probe frequencies the CAP trend closely follows the lower boundary of the ANF data cluster; but at 10 kHz, the CAP trend intersects the ANF trend.

Many studies report Q_{ERB} rather than Q_{10} values, the rationale being that the former provides a more robust measure because it incorporates all measurement points rather than only the points at the lowest threshold (CF) and 10 dB above. However, as pointed out in METHODS, this is an advantage only when sufficient data points are available around the nadir of the tuning curve. Fig. 15b compares the CAP-Q_{ERB} trend line (red, same curve as in Fig. 8b) with that from the ANF data. As in the case for the Q_{10}, the CAP-Q_{ERB} trend line is largely located within the cluster of ANF data points, but it is less congruent with the lower envelope of that distribution. This is in line with the earlier finding in Fig. 9b, where the course of the CAP-Q_{ERB} clearly differed from that of the ROEX prediction. To understand this discrepancy we studied the Q_{ERB}/Q_{10} relationships in cat for ANF and CAP. For ANFs, the Q_{ERB}/Q_{10} ratio is largely independent of species and CF (Shera and Guinan, 2003, footnote 6). Fig. 16 shows the Q_{ERB}/Q_{10} ratios derived from ANFs (gray dots and dashed line) and CAPs (blue line, data from Fig. 15) in cat.

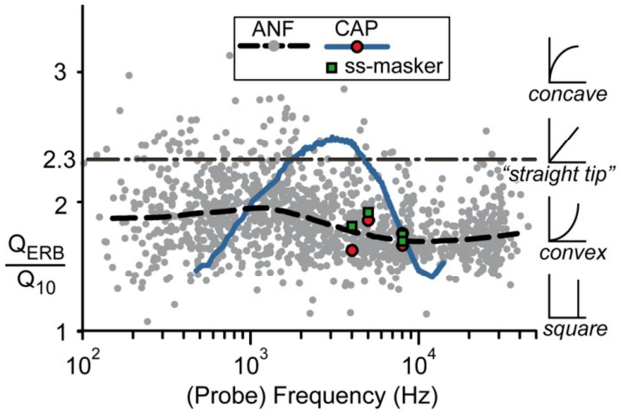


Figure 16. The Q ratio of the slopes near the nadir of the MTC as a function of probe frequency for cat, calculated as the ratios of Q_{ERB} and Q_{10} (blue solid line, data from trend lines in Fig. 8a and b). Ratios from AN fibers and a trend line are included as gray dots and the dashed line (trend obtained with *RLOWESS* function in MATLAB). The dashed dotted line indicates tuning curves with "straight" nadirs; above the line the nadir is concave, below, convex (see tuning curve cartoons at the right side of the figure).

To illustrate the ratio for different tip-curvatures, associated cartoons are shown at the right side of the figure. It is clear that the Q_{ERB}/Q_{10} ratio of the CAPs has a much larger dependence on frequency than the trend of the ANFs, which is rather stable with a mean value of 1.82. Our interpretation is that due to a small number of data points near the nadir of the MTCs the Q_{ERB} values for CAPs are unreliable. This is confirmed by the analysis of the MTCs from Fig. 14 which have much finer spaced notch-widths. The corresponding ratios of these double sided and single sided MTCs are included in Fig. 15 as the red circles and green squares, respectively: these values straddle the lower boundary of the ANF distribution of Q_{ERB} values rather than the mean trend (red curve). The associated Q_{ERB}/Q_{10} ratios depicted as the red circles and green squares in Fig. 16, also show more consistency with the ratios of ANFs. In summary, the CAP-Q_{10} values (Fig. 15a) are consistent between data obtained at high and at low resolution, while this is not the case for the Q_{ERB} values (Fig. 15b). This confirms our interpretation that Q_{ERB} factors are only reliable when enough data points are available around the nadir of the tuning curve. Because we do not have enough data with sufficient data points around the nadir, we restrict ourselves to Q_{10} factors in the remaining analysis.

As has been pointed out in the introduction, human sharpness of frequency tuning is unknown for single ANFs. CAP recordings provide an electrophysiological means to compare tuning across species, including humans. Moreover, CAP/ANF conversion functions of representative animals may allow us to estimate neural sharpness of tuning in animals or human where no measurements of single units are available. In order to estimate ANF tuning from CAP tuning in other species and human, we first have to assess the quantitative relationship between CAPs and single unit responses. In Fig. 17a the CAP-Q_{10} of chinchilla (red line) and cat (blue line) are compared with the trends of ANFs (chinchilla: from (Ruggero and Temchin 2005); cat: own data from Fig. 15). As in cat (see earlier discussion in *Comparison with single auditory nerve fibers*), the CAP-Q_{10} trend of chinchilla runs below the ANF trend, intersects it at 7.3 kHz and is slightly larger above this frequency.

The relationship between Q_{10} ratios of ANF and CAP for cat (blue) and chinchilla (red) are shown in Fig. 17b. Although the ratios in chinchilla are generally lower, the resemblance between the curves of the two animals is remarkable. Two regions can be distinguished that are separated by a deflection point at ~5 kHz. The first region, below the deflection point, shows a modest decrease with frequency. The second region has a much steeper declining slope and includes the intersection of the trend lines (cat: 10 kHz, chinchilla: 7.3 kHz), indicated by the dashed line.

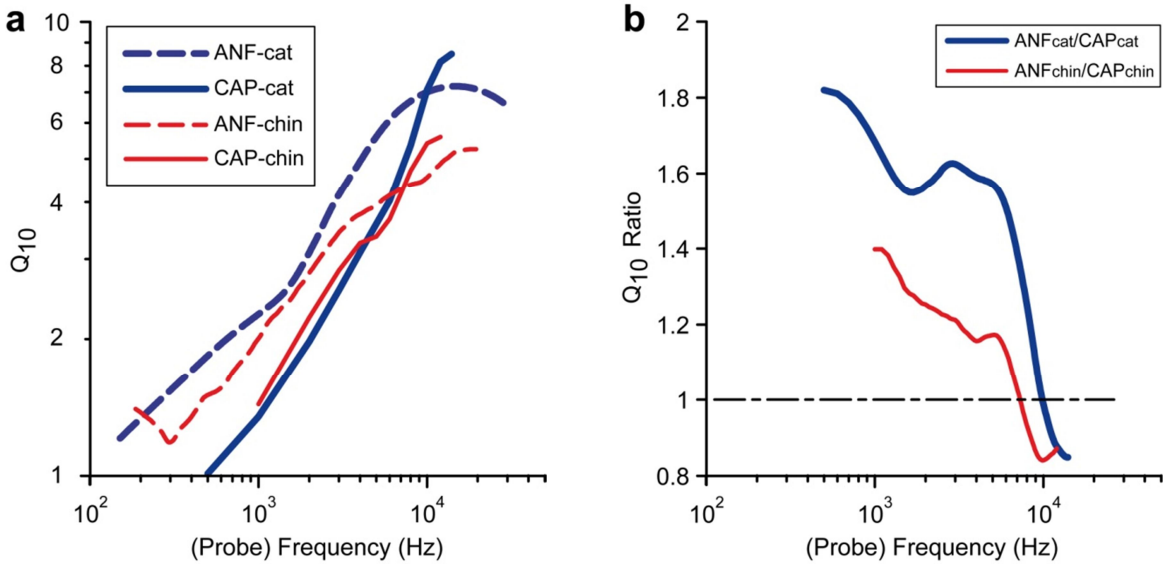


Figure 17. **a)** Comparison between Q_{10}-trend lines (trend obtained with *RLOWESS* function, MATLAB) from chinchilla and cat as a function of frequency. Cat: blue solid line: present CAP data, blue dashed line: ANF data (same data as in Fig. 15a); Chinchilla: red solid line: present CAP data, red dashed line: ANF data as presented in fig. 4 of (Ruggero and Temchin 2005). **b**) ANF-Q_{10}/CAP-Q_{10} ratios from cat and chinchilla data in panel **a**.

DISCUSSION

Whether frequency tuning is sharper in humans than in commonly studied laboratory animals is a matter of controversy (Ruggero et al., 1997; Shera et al., 2002; Oxenham and Simonson, 2006; Lopez-Poveda et al., 2007; Eustaquio-Martin and Lopez-Poveda, 2011; Joris et al., 2011). The aim of the present study was to examine whether the combination of forward masking and CAP recording provides a suitable means to investigate the basic frequency tuning of the neural input to the central nervous system. CAP recordings can be performed in humans in a minimally invasive way (Harrison et al., 1981), and may therefore provide comparative data across species. But the CAP is a mass potential, giving indirect information regarding tuning in single nerve fibers: "tricks" in stimuli and analyses are required to estimate frequency tuning. Behavioral studies also face this problem, and Oxenham and Shera (2003) argued that forward masking paradigms hold the greatest promise in this regard. However, from an extensive literature survey, Ruggero and Temchin (2005) concluded that "forward-masking psychophysical tuning curves […] greatly overestimate the sharpness of cochlear tuning in experimental animals and, hence, also probably in humans". To investigate the validity of forward-masked CAP responses as a probe recorded from two species for which abundant single nerve fiber data are available: cat and chinchilla.

We find that masking of CAP responses behaves as expected from cochlear and nerve physiology in a number of respects. The masking intensity curve to broadband noise at low probe level (Fig. 5, red curve) is very similar to that of tonal maskers in single fibers (Harris and Dallos, 1979). Also, recovery from masking as described by these authors applies in similar form to our data (Fig. 6). Masking tuning curves show a monotonic increase with notch-width, and they become narrower with increasing probe frequency (Fig. 7). Tuning sharpness, as quantified by Q_{10}, decreases for probe levels above ~45 dB SPL (Fig. 11). The slope of the growth of maskability with probe level is as expected from cochlear compression (Fig. 12 and Fig. 13). Single-sided and double-sided maskers result in quite similar estimates for Q_{10}, which validates the assumption of symmetry around the nadir of the tuning curve (Fig. 14 and Fig. 15). Finally and most importantly, CAP-Q_{10} values increase with probe frequency at a slope very similar to the trend in single ANFs (Fig. 15, Fig. 17 and Fig. 18), and the same basic relationship is found in the two species studied (cat and chinchilla), although the data in chinchilla showed more variability. We find that over most of the frequency range the CAP-Q_{10} closely follows the lower boundary of the distribution of neural Q_{10} values (Fig. 18), and underestimates the trend rather than overestimating it as seen in some of the available masking studies (symbols in Fig. 18). Only at high frequencies does the CAP trend cross the trend for single fibers: in cat this crossing point is near 9 kHz and in chinchilla it is near 7 kHz. Overall, the results show that the technique of notched-noise forward masked CAP measurements provides a reliable estimate of the lower boundary of frequency tuning in the auditory nerve. Even if it would be the case that psychophysically the paradigm results in an overestimate of frequency tuning (Ruggero and Temchin 2005), this is not the case for the paradigm used in this study.

In the present paper, we explore the paradigm as an empirical tool to obtain Q measures from CAP recordings. To fully understand the relationship between single fiber and CAP responses, it is necessary to 1) characterize the population response of fibers to the stimuli used, and 2) to characterize the "unit contribution" of each fiber at the round window. For example, the CAP response to click stimuli has been modeled convolving post-stimulus time histograms with spike-triggered averages of the round window potential for single fibers and summing across CF (Goldstein and Kiang, 1958; Antoli-Candela and Kiang, 1978; Wang, 1979; Prijs, 1986). Single fiber studies of forward masking are limited and restricted to probe frequencies at the CF of the fiber studied (Harris and Dallos, 1979): population data are needed to allow prediction of forward masked CAPs. In the absence of such data, auditory nerve models can be used to gain insight into the relationship between the notch-noise forward masked CAP data and single fiber data. A full treatment of factors generating the

patterns observed is deferred to a subsequent study and we here restrict ourselves to some general comments on potential mechanisms.

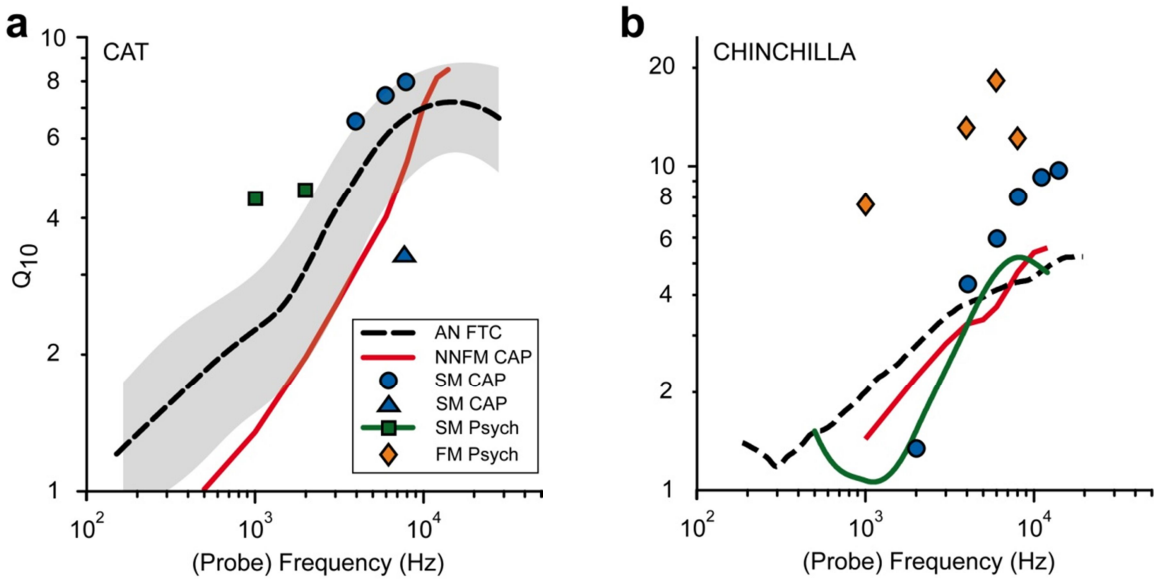


Figure 18. Comparison of sharpness between different paradigms and methods for cat (a) and chinchilla (b). The Q_{10} values from ANF and the NNFM paradigm for CAP are indicated by the dashed black and solid red lines, respectively; values from SM-CAP are indicated as blue circles and triangles. a: Q_{10} values from psychophysical studies using simultaneous masking are indicated by filled green squares (Pickles 1979, Fig. 7); SM-CAP: blue circles (Carlier et al. 1979) and triangles (Van Heusden and Smoorenburg 1981); the gray area indicates the upper and lower border of the AN-fiber's standard deviation. b: Q_{10} values from a psychophysical study using forward masking are indicated by yellow filled diamonds (McGee et al. 1976); trend from values of psychophysical studies using simultaneous masking are indicated by the green line (Salvi et al. 1982; Niemiec et al. 1992); SM-CAP: blue circles (Spagnoli and Saunders 1987).

Synchronization across nerve fibers is essential to obtain a measurable CAP (Goldstein and Kiang, 1958). Cochlear mechanisms cause a CF-dependent delay such that high-CF fibers tend to be more synchronized to each other than low-CF fibers. In Fig. 17b, we show the ratio between neural and CAP-Q_{10} values as a function of frequency, for cat (blue line) and chinchilla (red line). The trend is similar for the two species and shows two segments with different slope. We surmise that the general decrease in the ratio from low to high probe frequencies, as well as the accelerating trend in this decrease, reflects an increase in across-fiber synchronization with increasing probe frequency. However, it is difficult to estimate the interaction of probe frequency, notch bandwidth, probe and masker level, on response variables such as response amplitude and synchronization without further empirical data or use of a detailed nerve model.

A factor which has to be taken into account is the (a)symmetry of the excitation patterns. One of the assumptions made in the paradigm is symmetry in the excitation patterns around the probe frequency. Preliminary modeling suggests that the presence of asymmetry around the nadir of the MTC predicts lower Q values measured with (symmetrical) double-sided maskers. We made a limited number of measurements with single-sided maskers (Fig. 14), which have the ability to reveal asymmetry in MTCs, and indeed the Q factors yielded somewhat higher values than those obtained with double-sided maskers. This is illustrated in Fig. 15a where the Q_{10} values from the single-sided noise maskers (green squares) are consistently higher than the double-sided noise maskers (red circles).

Fig. 18 shows Q values from several sources (compiled by Ruggero and Temchin, 2005; for references see legend). Behavioral studies used simultaneous masking (Fig. 18a: green squares, Fig. 18b: green line) or forward masking (orange diamonds); CAP studies also used simultaneous masking (blue circles) or forward masking (red trend line: this study). The trend lines for single fiber tuning are also shown. Surprisingly, nearly all values based on simultaneous masking are above our CAP trend line (red) and even above the trend line for the neural data (dashed). As pointed out by Ruggero and Temchin (2005), the behavioral data obtained with forward masking (chinchilla) yield extremely high Q values compared to either the single-fiber nerve data or to our CAP data. Similar data and trends are shown for other species in Figs. 2 and 4 of Ruggero and Temchin (2005). The origin of the high Q values measured with behavioral forward masking in chinchilla (or in other species) is unclear. Ultimately, the behavioral data must reflect the interpretation by a central processor of activation patterns in the auditory nerve, and it is unclear which aspects of these patterns are critical (Relkin and Turner, 1988; Relkin and Smith, 1991). For example, “off-frequency listening” is mentioned by Ruggero and Temchin (2005) as a probable cause for the large difference between forward-masked behavioral and neural data. However, a possible peripheral factor leading to overestimation of sharpness is compression of the masker, also discussed by Eustaquio-Martin and Lopez-Poveda (2011). Note that compression can also occur in threshold tuning curves of auditory nerve fibers, especially for frequencies away from CF, which require higher signal levels to comply with the same response criterion. However, Q_{10} values for auditory nerve fibers are unlikely to be affected by compression, because the corresponding SPL values (at the tuning curve nadir and the range 10 dB above) are too low to be in the compressive range. The schematic input-output function in Fig. 13b shows two linear regions connected by a compressive region where larger masking levels are required to obtain the same effective masking as in the linear regions – an outcome which is phenomenologically similar to that of a high-Q system. Growth-of-maskability functions in

detection studies commonly show a slope ≥1 (Relkin and Turner, 1988; Relkin and Smith, 1991; Oxenham and Plack, 1997; Plack and Oxenham, 1998), as illustrated by the red lines (Fig. 13a). In contrast, in our data (Fig. 13a, green circles), these functions had a slope <1, suggesting that the probe was rather in the compressive region and the masker in the lower linear region, and *vice versa* in detection studies like most psychophysical studies. The difference may stem from the masking criterion. We surmise that the effective masking criterion, and therefore masking, is high in detection studies, while it is typically low and predefined in physiological studies (e.g., 33% here and in Dallos and Cheatham [1976], Fig. 6). Variables that increase masking (e.g., shortening of the masker-probe interval; increase in masker duration) can influence the masking criterion and thereby affect the estimated Q value. Most of our measurements were at low probe and masker levels: to the extent that the probe levels reached the compressive region of the input-output function, the Q values obtained with forward-masked CAPs would be underestimates of nerve-fiber tuning.

In psychophysical studies it is common practice to convert MTCs obtained with notched-noise maskers to tonal equivalent auditory filters using a power spectrum model of masking. In contrast to psychophysical studies, Q_{10} values extracted from auditory filters (Fig. 9) were, although not importantly, smaller than those directly obtained from MTCs. This suggests that, for cochlear tuning, masking with notched-noise maskers is similar to masking with tonal maskers. The lack of off-frequency listening and central auditory processing could explain the difference with psychophysical studies.

CONCLUSIONS

We investigated sharpness of CAP-tuning at the level of the auditory periphery using a notched-noise forward masking paradigm. The sharpness of tuning, quantified as a Q_{10}, showed relationships which were consistent with those of ANFs. The sharpness of tuning decreased with increasing probe level and increased with probe frequency. The obtained Q_{10} values were generally lower than those of ANFs and straddled the lower boundary of the distribution. The results support the use of forward-masked CAPs as an alternative measure to estimate ANF tuning and to compare frequency tuning across species.

ACKNOWLEDGMENTS

Supported by Research Fund KU Leuven (OT/09/50 and BIL11/12 and senior fellowship to Luis Robles). Damir Kovačić was supported by a Marie-Curie IEF Fellowship (GA 221755); Luis Robles was partly funded by Fondecyt 1080227. We thank Enrique A. Lopez-Poveda for valuable comments on a previous version of the manuscript and Almudena Eustaquio-Martin for providing us with the algorithm to fit the ROEX filter shapes to the experimental data.

REFERENCES

Antoli-Candela EJ, Kiang NYS (1978) Unit activity underlying the N1 potential. In: Evoked Electrical Activity in the Auditory Nervous System (Naunton R, Fernandez C, eds), pp 165-189. New York: Academic Press.

Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. Nature 451:197-201.

Carlier E, Lenoir M, Pujol R (1979) Development of cochlear frequency selectivity tested by compound action potential tuning curves. hear Res 1:197-201.

Dallos P (1973) The auditory periphery. New York: Academic Press.

Dallos P, Cheatham MA (1976) Compound action potential (AP) tuning curves. J Acoust Soc Am 59:591-597.

Eggermont JJ (1977) Compound actionpotential tuning curves in normal and pathological human ears. J Acoust Soc Am 62:1247-1251.

Eustaquio-Martin A, Lopez-Poveda EA (2011) Isoresponse versus isoinput estimates of cochlear filter tuning. J Assoc Res Otolaryngol 12:281-299.

Evans EF (1972) The frequency response and other properties of single fibres in the guinea-pig cochlear nerve. J Physiol 226:263-287.

Glasberg BR, Moore BC, Patterson RD, Nimmo-Smith I (1984) Dynamic range and asymmetry of the auditory filter. J Acoust Soc Am 76:419-427.

Goldstein MH, Kiang NYS (1958) Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. J Acoust Soc Am 30:107-114.

Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. J Neurophysiol 42:1083-1107.

Harrison RV, Aran JM, Erre JP (1981) AP tuning curves from normal and pathological human and guinea pig cochleas. J Acoust Soc Am 69:1374-1385.

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proc Natl Acad Sci U S A 108:17516-17520.

Kiang NYS, Moxon EC, Kahn AR (1976) The Relationship of Gross Potentials Recorded fron the Cochlea to Single Unit Activity in the Auditory Nerve. In: Electrocochleography (Ruben RJ, Elberling C, Salomon G, eds), pp 95-115. Baltimore: University Park Press.

Liberman MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am 63:442-455.

Lopez-Poveda EA, Barrios LF, Alves-Pinto A (2007) Psychophysical estimates of level-dependent best-frequency shifts in the apical region of the human basilar membrane. J Acoust Soc Am 121:3646-3654.

McGee T, Ryan A, Dallos P (1976) Psychophysical tuning curves of chinchillas. J Acoust Soc Am 60:1146-1150.

Nelson DA, Schroder AC, Wojtczak M (2001) A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 110:2045-2064.

Niemiec AJ, Yost WA, Shofner WP (1992) Behavioral measures of frequency selectivity in the chinchilla. J Acoust Soc Am 92:2636-2649.

Oxenham AJ, Plack CJ (1997) A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. J Acoust Soc Am 101:3666-3675.

Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J Assoc Res Otolaryngol 4:541-554.

Oxenham AJ, Simonson AM (2006) Level dependence of auditory filters in nonsimultaneous masking as a function of frequency. J Acoust Soc Am 119:444-453.

Patterson RD (1976) Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 59:640-654.

Patterson RD, Nimmo-Smith I, Weber DL, Milroy R (1982) The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold. J Acoust Soc Am 72:1788-1803.

Pickles JO (1979) Psychophysical frequency resolution in the cat as determined by simultaneous masking and its relation to auditory-nerve resolution. J Acoust Soc Am 66:1725-1732.

Plack CJ, Oxenham AJ (1998) Basilar-membrane nonlinearity and the growth of forward masking. J Acoust Soc Am 103:1598-1608.

Prijs VF (1986) Single-unit response at the round window of the guinea pig. Hear Res 21:127-133.

Relkin EM, Turner CW (1988) A reexamination of forward masking in the auditory nerve. J Acoust Soc Am 84:584-591.

Relkin EM, Smith RL (1991) Forward masking of the compound action potential: thresholds for the detection of the N1 peak. Hear Res 53:131-140.

Rhode WS, Roth GL, Recio-Spinoso A (2010) Response properties of cochlear nucleus neurons in monkeys. Hear Res 259:1-15.

Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. PhysiolRev 81:1305-1352.

Ruggero MA, Temchin AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. Proc Natl Acad Sci U S A 102:18614-18619.

Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L (1997) Basilar- membrane responses to tones at the base of the chincilla cochlea. J Acoust Soc Am 101:2151-2163.

Salvi RJ, Giraudi DM, Henderson D, Hamernik RP (1982) Detection of sinusoidally amplitude modulated noise by the chinchilla. J Acoust Soc Am 71:424-429.

Sayles M, Winter IM (2010) Equivalent-rectangular bandwidth of single units in the anaesthetized guinea-pig ventral cochlear nucleus. Hear Res 262:26-33.

Shera CA, Guinan JJ, Jr. (2003) Stimulus-frequency-emission group delay: a test of coherent reflection filtering and a window on cochlear tuning. J Acoust Soc Am 113:2762-2772.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci U S A 99:3318-3323.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2010) Otoacoustic estimation of cochlear tuning: validation in the chinchilla. J Assoc Res Otolaryngol 11:343-365.

Spagnoli SD, Saunders JC (1987) Threshold sensitivity and frequency selectivity measured with round window whole nerve action potentials in the awake, restrained chinchilla. Otolaryngol Head Neck Surg 96:99-105.

Temchin AN, Rich NC, Ruggero MA (2008) Threshold tuning curves of chinchilla auditory-nerve fibers. I. Dependence on characteristic frequency and relation to the magnitudes of cochlear vibrations. J Neurophysiol 100:2889-2898.

Van Heusden E, Smoorenburg GF (1981) Eighth-nerve action potentials evoked by tone bursts in cats before and after inducement of an acute noise trauma. Hear Res 5:1-23.

Wang B (1979) The relation between the compound action potential and unit discharges of the auditory nerve. In, p 219: Massachusetts Institute of Technology.

Chapter 3

Electrophysiological evidence for sharp cochlear tuning in humans

ABSTRACT

Cochlear frequency tuning is fundamental to hearing. The sharpness of tuning in humans is unknown and is generally thought to be similar to that in the best studied laboratory mammals, but behavioral and acoustic measurements suggest sharper tuning in humans. We developed a procedure to estimate neural frequency tuning through potentials measured in the middle ear of healthy volunteers, and found sharper tuning than in commonly studied mammals.

INTRODUCTION

Frequency selectivity is fundamental to hearing. It refers to the ability to resolve different spectral components in complex sounds. This ability first arises in the cochlea and is pervasive in behavioral and neural responses. Frequency selectivity is classically assessed by measuring the sharpness of (frequency) tuning, expressed as a quality (Q) factor (see Methods), and has been studied most extensively in single auditory-nerve fibers, which innervate the inner hair cells along the frequency selective cochlear basilar membrane (Robles and Ruggero, 2001). Such recordings require penetrating microelectrodes into the exposed auditory nerve and are not possible in humans. Measurement of auditory-nerve frequency tuning in human thus requires indirect methods. Despite some early qualitative data from patients suggesting sharp tuning (Harrison et al., 1981), it is generally thought that human frequency selectivity is similar to that in commonly studied laboratory mammals (Glasberg and Moore, 1990; Ruggero and Temchin, 2005). Two recent studies, using stimulus frequency otoacoustic emissions (SFOAEs) (Shera et al., 2002) and behavioral measurements (Oxenham and Shera, 2003), have challenged this view, reporting sharper tuning in humans. It has been argued that these results reflect experimental artifacts rather than real species differences (Ruggero and Temchin, 2005; Eustaquio-Martin and Lopez-Poveda, 2011).

To obtain a more direct, neural measure, we revisited and improved a method used in early clinical studies (Portmann and Aran, 1971; Gibson et al., 1977), and adapted it for use in normal hearing volunteers. An electrode is punctured through the anesthetized eardrum to record sound-evoked potentials from the cochlear bony capsule in the middle-ear cavity, from which compound action potentials (CAPs) – the summed responses of synchronized auditory nerve fibers – are extracted. To achieve stable recordings over several hours, we developed a closed system allowing transtympanic electrode placement under visual control combined with *in situ* acoustic calibration (Methods).

CAPs reflect activity of nerve fibers tuned to different frequencies (Goldstein and Kiang, 1958; Antoli-Candela and Kiang, 1978), so that special stimulus paradigms are needed to assess frequency selectivity. We used forward masking (Harris, 1979), where a pure tone (the probe), fixed in level and frequency, is used to probe the imprint of the spectrally manipulated preceding signal (the masker). To facilitate comparison with behavioral studies, we applied in human and macaque monkey a notched-noise forward masking (NNFM) paradigm (Patterson, 1976; Oxenham and Shera, 2003), which we previously validated in cat and chinchilla (Verschooten et al., 2012).

METHODS

All procedures, for both animal and human experiments, were approved by the Medical Ethics Committee and Animal Ethics Committee of the University of Leuven and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Human subjects. Volunteers were recruited from the local student population with advertisements and were between 20 and 35 years old; all gave written informed consent. They were requested to avoid exposure to loud sounds such as rock concerts in the days preceding the recording session. The day before, or morning of, the session, the subject's hearing was screened including an anamnesis for hearing problems, a pure tone audiogram (thresholds had to be < 20 dB HL, 125 Hz - 8 kHz), tympanometry to assess middle ear function, and an otoscopy by an ENT doctor. The duration of the experimental session varied between 1 and 4 hours; subjects could stop the session at any time. To minimize electrical and acoustical interference, all experiments, including these with monkey, were conducted in a double-walled soundproofed and faradized booth (Industrial Acoustics Company, Niederkrüchten, Germany). Subjects chose a comfortable reclined position on a bed and were asked to remain still or to nap during the recordings. Subjects and experimentalists, while in the sound booth, were grounded to the booth via an antistatic wrist strap. An observer was with the subject at all times and was in visual communication with the experimenters who controlled the experiment from outside the booth. A total of 18 subjects (15 female, 3 male) was studied. In 7 subjects, recordings were not possible for a variety of reasons (muscle or movement artifacts, poor signal-to-noise ratio, technical problems, anatomical factors preventing safe placement of the electrodes). In 2 of the 12 remaining subjects, some measurements were obtained but were of insufficient quantity or quality.

Monkeys. Recordings were obtained from one ear in four monkeys (Macaca mulatta) which were also used in chronic visual experiments (an 8.9 kg adult male, a 4.8 kg juvenile male; and two juvenile females of a 6.3 kg and 4.7 kg). Prior to the experiments, dissections on formaline-preserved temporal bones were performed to study the desired trajectory and placement of the needle electrode. The recording procedure in monkey was similar to that in human, with the main difference being the presence of anesthesia. Induction of general anesthesia was done with a mixture of ketamine (3 mg/kg) and medetomidine (0.050 mg/kg), i.m. The same mixture was administered i.v. for maintenance, through a venous cannula inserted for administration of lactated Ringer's solution. The tympanic membrane was checked otoscopically. A simple custom ear mold was made. The needle electrode was placed under guidance of a surgical microscope and was fixed relative to the head with a mechanical manipulator. The duration of the total experimental session including the placing of the electrode was between 4 and 6 hours. After the experiments, atipamezole (0.2 mg/kg) was administered to reverse the sedative effect of medetomidine; after awakening of the animal, it was observed for a certain period of time.

Trans-tympanic Procedure. A trans-tympanic procedure was developed, extensively tested, and practiced on 20 fresh human cadavers in the university hospital. For every subject, a custom silicone ear mold (Dentsply, Aquasil Ultra XLV regular) was made to optimize low frequency performance of the earphone-speaker and for acoustical reproducibility throughout the procedure. The ear mold contained two casted openings to hold tubes of 2 mm diameter for needle insertion, visualization, acoustic stimulation, and calibration. The complete acoustic system was calibrated *in situ*, with a probe-microphone (Etymotic Research, ER-7C) close to the tympanic membrane. The earphone-speaker was connected to one of the openings of the ear mold via a plastic T-piece which also served as access port for a rigid endoscope with camera (R. WOLF, 8654.402 25 degree PANOVIEW; ILO electronic GmbH, XE50-eco X-TFT-USB) to visualize the ear canal and tympanic membrane. During the acoustic calibration all openings were sealed with Audalin acrylic impression compound (Microsonic); a tiny opening in one of the tubes prevented static pressure build-up. Before the needle-electrode was inserted, the tympanic membrane and ear canal were locally anesthetized with Bonain's solution (equal amounts of cocaine hydrochloride, phenol and menthol), which was aspirated after about 30 minutes. A short sterile plastic tube was inserted in the mold to accommodate the sterile needle-electrode. Ground and reference electrodes were connected to the equipment. The needle-electrode (TECA, sterile monopolar disposable, 75 mm x 26 G, 902-DMG75-TP), was inserted and placed through the tympanic membrane on the cochlear

promontory or in the niche of the round window under visual endoscopic control. To maintain its position and to ensure good electrical contact, the needle-electrode was maintained under slight tension with rubber bands supported by a frame, which was positioned above the external ear with soft velcro straps around the head. Subjects usually had a short-lasting and vague sensation of touch during insertion of the electrode. The openings of the tubes were sealed with Audalin and the needle-electrode was connected to the preamplifier. Usually the subject's right ear was studied: there was no experimental manipulation of the other ear. The session was terminated after 4 hours, or earlier when the recordings became too noisy or when the subject expressed the desire to stop. The needle electrode was then simply pulled back and the ear mold was removed. The session was concluded with an otoscopic examination. In no cases was there an eardrum perforation larger than expected from the needle's diameter (0.46 mm). Subjects were requested not to shower or swim in the week following the recording session. An ENT doctor was available during the week after the experiment to address any worries.

Acoustical stimulation. Stimuli were generated with custom software and a digital sound system (Tucker-Davis Technologies, system 2, sample rate: 125 kHz/channel) consisting of a digital-to-analog converter (PD1), a digitally controlled analog attenuator (PA5), a headphone driver (HB7) and an electromagnetically shielded earphone-speaker (Etymotic Research, ER1 or ER2, 20 Hz - 16 kHz). The transducer was connected with plastic tubing to the ear mold.

Electrophysiological recording and analysis. Auditory evoked potentials were measured using a low noise differential preamplifier (Stanford Research Systems, SR560). The signal input was connected to the needle-electrode; the reference input was connected to an earlobe clamp (with conductive gel) and the ground input was connected to a standard disposable surface electrode placed at the mastoid. All contacts were made on the side ipsilateral to the recording. The battery-operated preamplifier was galvanically isolated (A-M systems, Analog stimulus isolator Model 2200) from the mains-powered equipment. Before the signal was recorded (TDT, RX8, ~100 kHz/channel, max. SNR 96 dB), stored and analyzed (Matlab), the signal was further amplified (DAGAN, BVC-700A) and band pass filtered (100 Hz - 10 kHz, cut-off slopes 12 dB/octave). All relevant signals (the recorded signal and stimulus and synchronization pulses) were visualized (LeCroy, WaveSurfer 24Xs) and monitored during the session. The CAP-response was de-noised with a wavelet filter. The measured responses contained significant uncorrelated background noise, which was reduced

by averaging multiple repetitions (between 128 and 1024, depending on the background noise level). The CAP amplitude was measured between the first negative trough (N1) and first positive peak (P1, Fig. 1). If P1 was not clearly defined, the amplitude was measured between N1 and the second positive peak (P2), or more generally, the positive maximum.

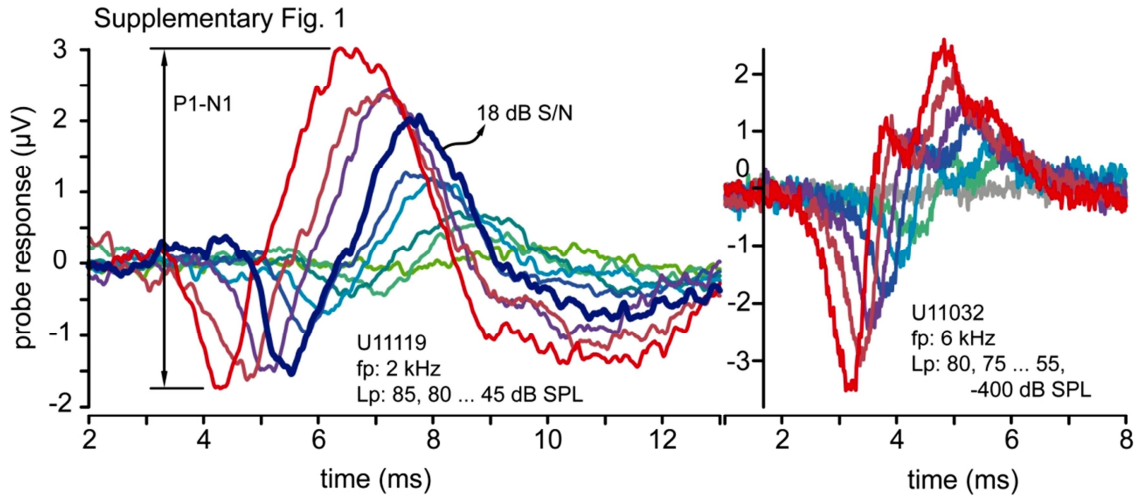


Figure 1. CAP response as a function of probe level (in dB SPL) for two different probe frequencies. Left panel for a probe frequency (f_P) of 2 kHz; right panel for 6 kHz. The magnitude of the CAP is obtained as the maximum voltage and is indicated by the arrow line. The response with the thick blue line in the left panel is an example of a response that meets our signal to noise ratio (S/N) criterion of 18 dB. The indication -400 dB SPL means no masker present. Data in left and right panel are not de-noised and are derived from 2 different subjects.

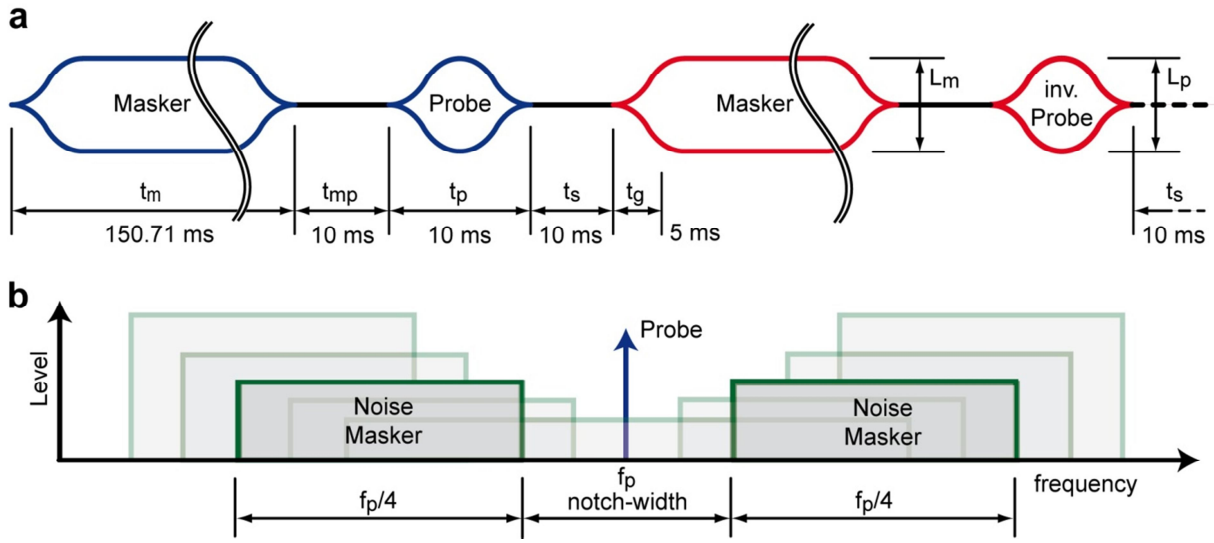


Figure 2. Schematic representation of the stimulus. (**a**) time domain, (**b**) spectral domain.

Stimulus. Schematic representations of the stimulus in the time and frequency domain are shown in Fig. 2a. A notched-noise masker of variable level is followed by a 10 ms tonal probe at a fixed predetermined level. To reduce transient spectral splatter, probe and masker were gated with 5 ms raised-cosine ramps.

To eliminate the cochlear microphonic (contribution from the hair cells) from the averaged waveform, the stimuli were alternated in polarity (Dallos, 1973) in successive presentations. The masker duration was set to 150.71 ms. To avoid overlap between the off-set and on-set responses (CAP) of masker and probe, silent intervals t_{mp} and t_s of 10 ms were included. Fig. 2b depicts the idealized frequency spectrum of the stimulus. Spectrally, two noise bands with equal bandwidth ($f_p/4$) straddle the probe frequency (f_p), so as to create a notch symmetrically spaced around f_p. The masker notch-width and level are the main experimental variables. The masker level is expressed as an overall level (total noise band in dB SPL).

Measurement procedure.

1) *Fixed probe level*. To find a suitable fixed probe level, the response to many probe stimulus levels was first measured in the absence of a masker (Fig. 1). The probe level yielding an 18 dB SNR was determined and used in the remaining procedure.

2) *Masking criterion and masker reference level*. The probe tone was delivered at the fixed probe level determined in the previous step, and a (no-notch) broadband masker was presented at several levels (Fig. 3a) to generate a masking curve (Fig. 3b). A coarse tracking procedure with fitting was used to search the "masker reference level" i.e. the masker level that suppressed the CAP-response by the masking criterion (33% in Fig. 3b).

3) Q_{10}. The masker was then presented at 10 dB above the masker reference level and varied in notch-width to determine the width resulting in the same amount of masking of the CAP-response (33%) as the (no-notch) broadband noise masker. The sharpness of tuning, here expressed as a $Q_{10,}$ was obtained as the probe frequency divided by this notch-width (Fig. 4).

The average measurement time needed for determination of one Q value was minimally 1 hour. Previous work (Verschooten et al., 2012) shows that the above method gives Q-values similar to those obtained with a ROEX filter.

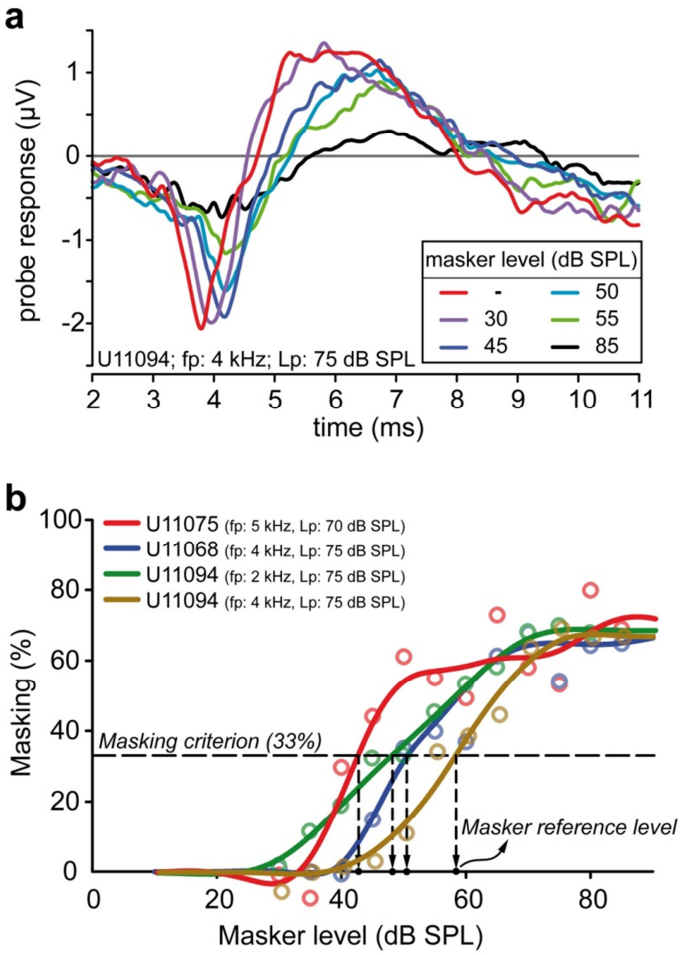


Figure 3. CAP masking as a function of masker level (L_m); the masker is a continuous broadband noise. (a) Example of masked CAP responses (not de-noised). (b) Masking curves for different probe frequencies (f_P). The dashed line indicates the masking criterion used in the paradigm. The arrow lines indicate the corresponding masker reference levels for different subjects and different probe frequencies (L_P).

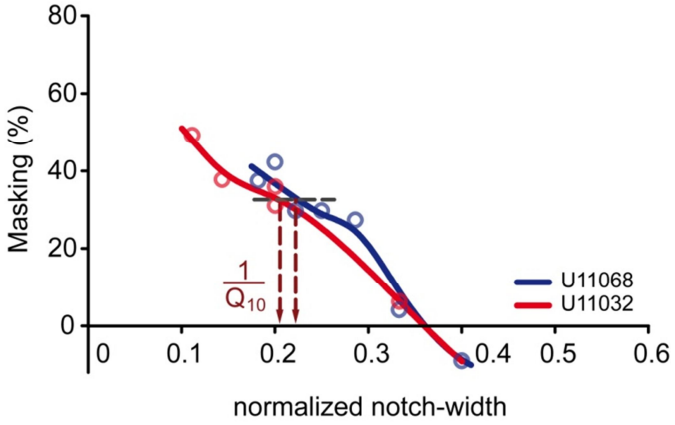


Figure 4. Masking as a function of normalized notch-width. Data are shown for 2 subjects. Probe frequency was 4 kHz; masker levels were fixed at 10 dB above the masker reference level. Horizontal dashed line indicates the masking criterion, here 33%. Vertical lines and arrows indicate the notch width giving the Q_{10}.

Stimulus levels. CAPs measured at the bony capsule of the cochlea in humans and monkeys were smaller and noisier than in cat and chinchilla. To comply with the predetermined SNR (~18 dB), stimulus levels in human, relative to cat and chinchilla, needed to be approximately 30 dB higher for both the probe (human: 55-75 dB SPL; monkey: 75-80 dB SPL) and the masker (human: 50-70 dB SPL; monkey: 40-65 dB SPL). It is well-known that cochlear frequency selectivity decreases with level. This has been extensively documented for iso-input measurements (Robles and Ruggero, 2001). Modeling studies suggest that for iso-response measurements, as used here, Q-values can actually increase with level (Eustaquio-Martin and Lopez-Poveda, 2011) so that we need to consider the possibility that higher Q values in human are due to the difference in stimulus levels. However, in the NNFM CAP study in animals (Verschooten et al., 2012) such an increase in sharpness with probe level was not observed. In the contrary, for probe levels below ~50 dB SPL there was almost no dependence on Q_{10}, and above this level the Q_{10} was negatively correlated with probe level (Verschooten et al., 2012). Consequently, if higher stimulus levels would affect our Q estimates, they would rather be underestimated than overestimation.

RESULTS

We first determined a probe level giving a CAP response with a signal-to-noise ratio (SNR) of ~18 dB (Fig. 1). Second, a broadband noise masker (without spectral notch) was added, the level of which was adjusted to cause a 33% CAP reduction (masking criterion) (Fig. 3). A notch was then inserted in the masker: by varying the notch-width and the level of the masker while keeping the CAP response constant (at 33% reduction), a masking tuning curve can be constructed (Verschooten et al., 2012). The restricted measurement time available in human subjects did not allow measurement of full tuning curves, so we searched directly for the 10 dB bandwidth, i.e. the notch-width requiring a 10 dB increase in masker level (Fig. 4), and calculated Q_{10} (probe frequency/bandwidth).

The CAP-Q_{10} values in human (Fig. 5a) and in monkey (Fig. 5b), obtained from 9 subjects are consistent with those previously obtained in animals (Verschooten et al., 2012), in several respects. First, the CAP-Q_{10} shows a monotonic positive dependence on probe frequency, indicating that sharpness of tuning increases with frequency. Second, the relative standard deviation (C_v, coefficient of variation; calculated as the estimated standard deviation divided by the sample mean) of the mean of the population in human and monkey, on average

over all measured frequencies (human: C_v = 5.5%, 16.7%, 9.9% [min., max., average across all frequencies]; monkey: C_v = 11.9%, 25.9%, 18.5%), are comparable to that in cat (C_v = 5.7%, 24.0%, 13.2% [min., max., avg], calculated from Fig. 8a in Verschooten et al., 2012). The coefficients of variation within individual subjects, also on average over all measured frequencies (human: C_v = 10.1%, 19.5%, 15.2% [min., max., avg]; monkey: C_v = 4.6%, 9.2%, 7.1%), is slightly larger for human than for macaque and is probably due to muscle artifacts (head movements, eye blinking, etc.) in awake subjects. Moreover, in some restless human subjects (n=2) the variance was very large and were in this cases clearly due to excessive head movements; these experiments were excluded from the data.

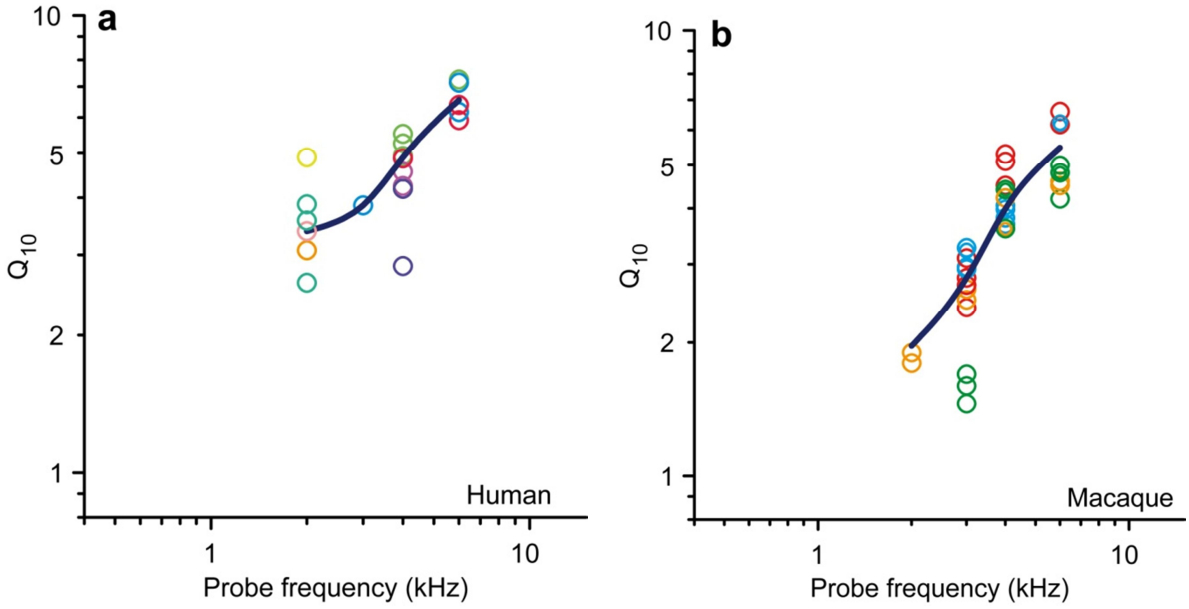


Figure 5. Sharpness of frequency tuning assessed with compound action potentials. CAP-Q_{10} as a function of the probe-frequency using the NNFM-paradigm. a) Obtained in 9 normal hearing humans. b) Obtained in 4 macaque monkeys. Different subjects are indicated by different colors. The black lines are trend line through the data points (Robust- LOESS function, MATLAB).

We compared the CAP-Q_{10} trends for the 4 species (Fig. 6): on average the tuning for human is 1.59, 1.56 and 1.31 times sharper than in cat, chinchilla and monkey, respectively. Remark that the CAP-Q_{10} trend of monkey for frequencies between 2 and 4 kHz is significant lower than that for frequencies between 4 and 6 kHz.

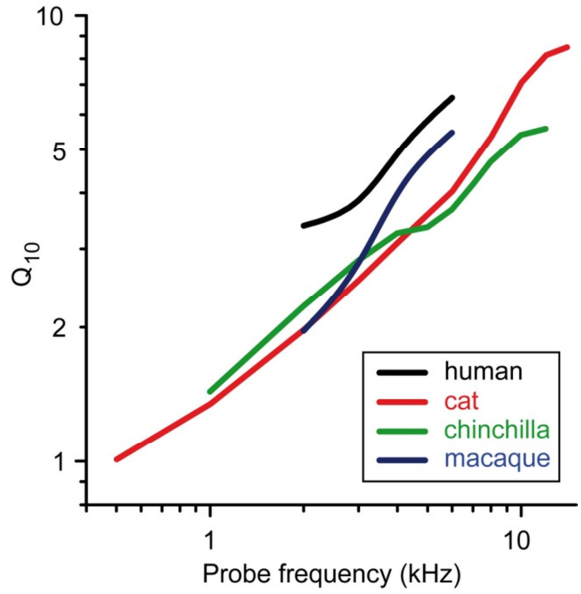
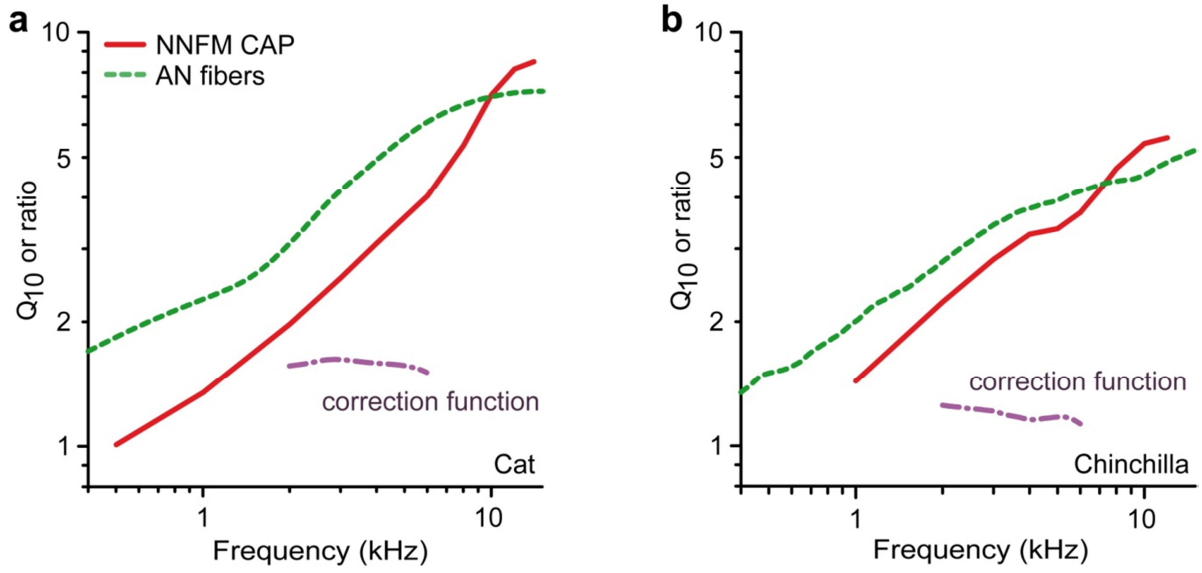


Figure 6. Comparison of CAP-Q_{10} trend lines in human, cat (Verschooten et al., 2012), chinchilla (Verschooten et al., 2012) and monkey. Human and monkey data are redrawn from Fig. 5.

To estimate the frequency tuning for single fibers in humans, we make the assumption that the relationship between CAP and Q_{10} values from single AN fiber is more or less similar across species. For cat, chinchilla and monkey, Q_{10} values are available for both single AN fibers and CAP potentials. From these sets of data, we calculated conversion functions (dashed-dotted lines in Fig. 7, respectively) over the frequency range of our CAP measurements in humans and macaques (2 to 6 kHz).



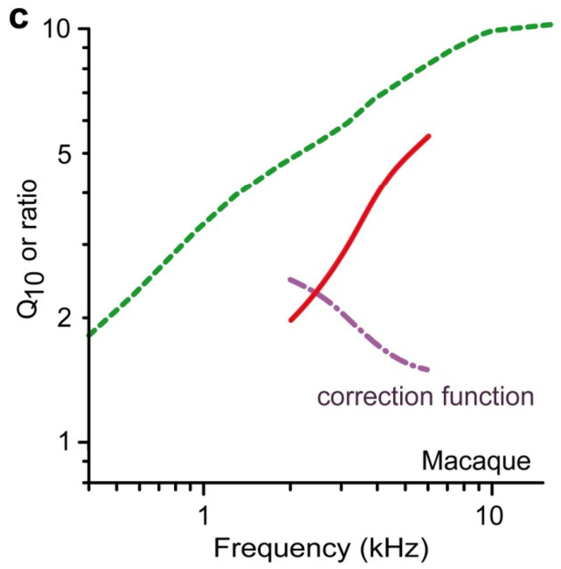


Figure 7. Conversion functions between sharpness of tuning (Q_{10}) of AN fiber data and CAP data. The purple dashed dotted lines are the conversion functions between the AN fiber data (green dashed lines) and CAP data (red lines; copied from Verschooten et al. [2012] and Fig. 6). The frequencies are limited to the measured frequency range in human (2 to 6 kHz). (**a**) For cat ($\overline{x}$ = 1.6). The AN fiber trend is derived from unpublished data from our laboratory. (**b**) For chinchilla ($\overline{x}$ = 1.2). The AN-fiber trend is redrawn from Figure 2B in Ruggero and Temchin (2005). (**c**) For macaque monkey ($\overline{x}$ = 1.8). AN fiber data from (Joris et al., 2011).

Multiplication of our human CAP-Q_{10} trend (black line, Fig. 6) with the conversion functions of the different species in Fig. 7 gives an average estimate (black line, Fig. 8) for the tuning of single AN fibers in human.

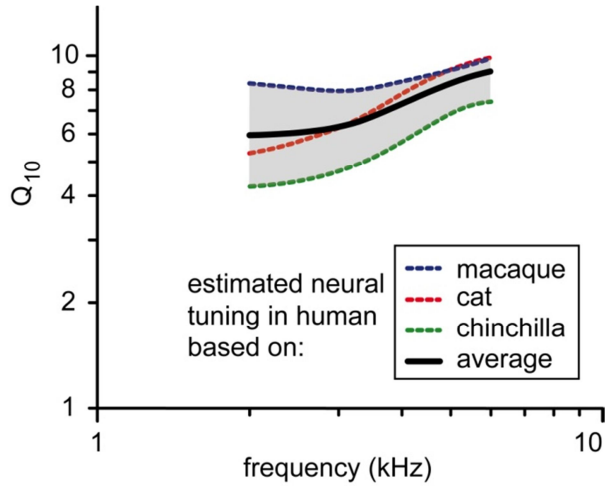


Figure 8. Estimated neural tuning in human obtained by the conversion of the human CAP-Q_{10} trend line (black line, Fig. 6) using animal (cat, chinchilla and monkey) conversion functions in Fig. 7. The average of these estimates is indicated by the black line. The range between the upper and lower estimates is indicated by the gray area.

Under the assumption that the (unknown) conversion function for humans would be located between the upper and lower estimated boundaries (gray area, Fig. 8), several observations can be made. First, the estimated neural tuning in human indicated by the gray area in Fig. 9 (same as that in Fig. 8) is clearly above the trends of single AN fiber tuning measured in a number of experimental animals (indicated by the shaded area; macaque not included). Second, this gray area surrounds the trend line reported for single AN fibers in macaque monkeys (gray line; from Joris et al., 2011). In addition, the average estimated

human tuning (black line) suggests that tuning in human is a little larger, but comparable, to that in macaque. Third, above 4 kHz the average estimated sharpness is similar to that reported using non-electrophysiological techniques (dotted line: SFOAEs from Shera and Guinan, 2003; dashed line: psychophysics from Oxenham and Shera, 2003); below 4 kHz it is broader, though not for the upper estimate of the gray area that rather coincides with the non-electrophysiological results. Remark that the upper bound is mainly determined by the conversion function of macaque. This conversion function is not flat, such as for the other species, but is significant larger below 4 kHz due to a lower CAP-Q_{10} trend with respect to single AN fibers (Fig. 7c) and also to the other species (Fig. 6).

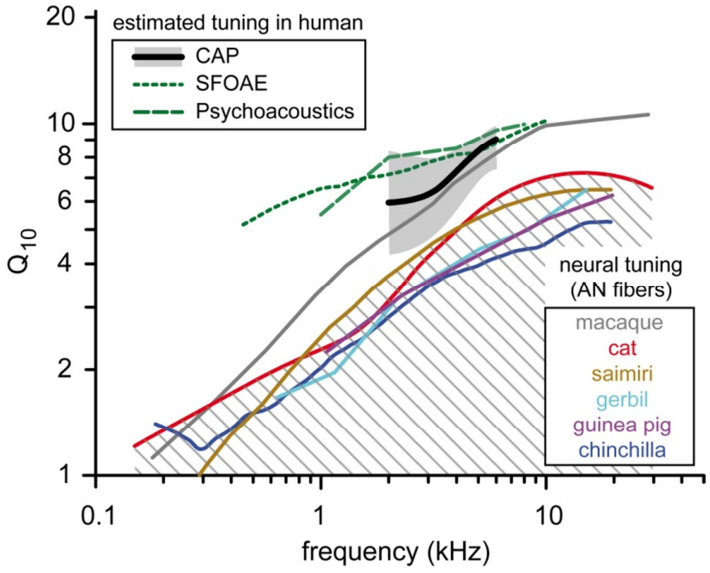


Figure 9. Comparison amongst estimated sharpness of neural tuning in human (black line) with that of single AN fibers in different species (cat: archival data from our laboratory, macaque: Joris et al., 2011; remaining trends redrawn from Ruggero and Temchin, 2005) and with that of other indirect measures of human frequency tuning (dotted line: SFOAE from Shera et al., 2002; dashed line: psychophysics from Oxenham and Shera, 2003; both are converted from Q_{ERB} with conversion factor 0.52). The range of the estimated neural tuning in humans is indicated by the gray area. The outline of the tuning of AN fibers, except for macaque, is indicated by the shaded area.

CONCLUSIONS

In this study we obtained electrophysiological estimates of peripheral frequency tuning in normal hearing volunteers. The results showed sharper frequency CAP-tuning in humans than in cat, chinchilla and most probably also in monkey. Furthermore, the good agreement between the estimated human neural sharpness and the single nerve fiber data of macaque monkey suggests that this sharp tuning is a feature of old-world primates.

ACKNOWLEDGMENTS

Supported by the Fund for Scientific Research – Flanders (G.0633.07) and Research Fund KU Leuven (OT/09/50). We thank R. Engelen and E. Kusseneers for assistance with the experiments.

REFERENCES

Antoli-Candela EJ, Kiang NYS (1978) Unit activity underlying the N1 potential. In: Evoked Electrical Activity in the Auditory Nervous System (Naunton R, Fernandez C, eds), pp 165-189. New York: Academic Press.

Dallos P (1973) The auditory periphery. New York: Academic Press.

Eustaquio-Martin A, Lopez-Poveda EA (2011) Isoresponse versus isoinput estimates of cochlear filter tuning. J Assoc Res Otolaryngol 12:281-299.

Glasberg BR, Moore BC (1990) Derivation of auditory filter shapes from notched-noise data. Hear Res 47:103-138.

Gibson WPR, Moffat DA, Ramsden RT (1977) Clinical electrocochleography in the diagnosis and management of Ménière's disorder. Audiology 16:389-401.

Goldstein MH, Kiang NYS (1958) Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. The Journal of the Acoustical Society of America 30:107-114.

Harris DM (1979) Action potential suppresion, tuning curves and thresholds: comparison with single fiber data. Hear Res 1:133-154.

Harrison RV, Aran JM, Erre JP (1981) AP tuning curves from normal and pathological human and guinea pig cochleas. Journal of the Acoustical Society of America 69:1374-1385.

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proceedings of the National Academy of Sciences of the United States of America 108:17516-17520.

Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J Assoc Res Otolaryngol 4:541-554.

Patterson RD (1976) Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 59:640-654.

Portmann M, Aran JM (1971) Electrocochleography. Laryngoscope 81:899-910.

Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. Physiological reviews 81:1305-1352.

Rose JE, Hind JE, Anderson DJ, Brugge JF (1971) Some Effects of Stimulus Intensity on Response of Auditory Nerve Fibers in the Squirrel Monkey. Journal of Neurophysiology 34:685-699.

Ruggero MA, Temchin AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. Proceedings of the National Academy of Sciences of the United States of America 102:18614-18619.

Shera CA, Guinan JJ, Jr. (2003) Stimulus-frequency-emission group delay: a test of coherent reflection filtering and a window on cochlear tuning. Journal of the Acoustical Society of America 113:2762-2772.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proceedings of the National Academy of Sciences of the United States of America 99:3318-3323.

Verschooten E, Robles L, Kovacic D, Joris PX (2012) Auditory nerve frequency tuning measured with forward-masked compound action potentials. Journal of the Association for Research in Otolaryngology : JARO 13:799-817.

Chapter 4

Upper frequency limit of phase-locking in stimulus evoked potentials

ABSTRACT

The role of the fine-structure in the perception of sounds has been debated for more than a century. It is clear from binaural behavioral studies that neural coding of fine-structure is present in humans at frequencies <1.5 kHz. However, the upper limit of this coding is contentious and is thought by many to extend up to 8-10 kHz. Of the presently available techniques, only the recording of mass neural potentials offers the promise to provide a physiological estimate of neural phase-locking in humans. A challenge is to disambiguate neural phase-locking from the receptor potentials. We studied mass potentials recorded on the cochlea and auditory nerve, and used several experimental manipulations to isolated the neural contribution in these potentials. We find a surprisingly large neural contribution in the signal recorded on the cochlea. The results suggest that this is a promising approach to examine neural phase-locking in humans.

INTRODUCTION

Stimulus related neural synchronization is a fundamental property generated in the mammalian cochlea. It refers to the ability of single auditory nerve fibers to represent temporal aspects of the stimulus waveform. In response to a low frequency tone, the probability of discharge is maximal at a preferred phase angle in the cycle of the tone. This type of neural synchronization to the fine-structure of the stimulus waveform is often called phase-locking and has been proposed to be involved in, or even to be critical to, many attributes of auditory perception (reviewed in Moore, 2008) and to underly hearing disorders (Kraus et al., 2000; Starr et al., 2001). Phase-locking in single fibers declines above a certain cutoff frequency and becomes undetectable above an upper frequency limit (Rose et al., 1967; Rose et al., 1968; Johnson, 1980; Joris et al., 1994a), which differs between species. Phase-locking to tones has been most extensively studied in the cat (Johnson, 1980; Rhode and Smith, 1985; Blackburn, 1992; Joris et al. 1994b) using the vector strength metric (Goldberg and Brown, 1969): it decreases at a cut-off frequency of about 1 kHz and becomes insignificant above ~5 kHz.

The cut-off frequency and upper frequency limit of phase-locking in humans are unknown. Psychophysical studies show a hard limit of <1.5 kHz up to which the binaural system can make use of stimulus fine-structure (Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Schiano et al., 1986; Dunai and Hartmann, 2011). This suggests that the limit of phase-locking to fine structure in human may be lower than in laboratory animals like the

cat (Joris and Verschooten, 2013), but it has also been argued that the limit is actually higher in humans than in experimental animals (Heinz et al., 2001; Moore, 2008). The technique used in animals to quantify phase-locking in single auditory nerve fibers cannot be applied in humans, because it requires intracranial access and invasive penetrating microelectrodes into the auditory nerve (AN).

Mass potentials measured at the AN, called the auditory nerve neurophonic, are thought to reflect synchronous activity of auditory nerve-fibers (Snyder and Schreiner, 1984). Such potentials provide a possible means to estimate and compare limits of phase-locking across species. Unfortunately, measuring such potentials on the human AN is not a matter of course because of the invasiveness. Studies in animals (Henry, 1995; He et al., 2012; Lichtenhan et al., 2013) reported the presence of a phase-locked neural component in the electrical response at the cochlear round window (RW); we refer to this as the round-window neurophonic (RW-NP). It is possible to record from the human RW using a minimally invasive approach through the middle ear (Eggermont, 1976a). Thus, if present in humans, the RW-NP would provide a means to physiologically assess human neural phase-locking. However, a problem arises because the potential at the RW, traditionally referred to as the cochlear microphonic (CM), is thought to be dominated by an ensemble potential that originates from the receptors. Classic studies of CM established that it is generated by the transduction currents of hair cells (Dallos, 1973). The extent to which AC potentials recorded at the RW reflect a contribution from the hair cells (henceforth referred to as the RW-CM) or from neural phase-locking (RW-NP) has not been characterized. It is well-documented that neural phase-locking shows a steep roll-off above ~1 kHz (in cat), while studies of the CM have shown an AC component extending to much higher frequencies. The entanglement of the potentials derived from the receptors and the auditory nerve therefore necessitates a careful separation of these two contributions.

The nerve shows two properties that may allow disambiguation of its contributions, relative to the receptor potentials, to the potential measured at the RW: rectification and adaptation. Neural rectification introduces higher harmonics, which are reported to be very small in the CM. The even harmonics, which are dominated by the 2nd harmonic, are easily extractable and have recently been proposed as an audiological tool (Lichtenhan et al., 2013). The other property, adaptation, causes temporary suppression of neural activity in the presence of a sufficiently strong preceding signal which is called forward masking. Forward masking of auditory neurophonic was demonstrated at the AN (Snyder and Schreiner, 1985) and RW (Henry, 1995).

Here, we develop a method, in a species for which phase-locking at the level of single AN fibers is well-characterized, to characterize neural phase-locking based on RW measurements. A combination of forward masking with stimulus polarity reversals was used to quantify RW-NP and RW-CM in cat. Furthermore, simultaneous recordings at the AN and administration of TTX at the RW were used to validate the method.

METHODS

Surgical Preparation

The experiments were conducted using 4 adult cats. All procedures were approved by the KU Leuven Ethics Committee for Animal Experiments and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Animals were briefly examined for clear external and middle ears. Cats were anesthetized with a mixture of ketamine (20 mg/kg) and acepromazine (0.2 mg/kg) administered intramuscularly. A venous cannula was then placed to infuse lactated Ringer's solution and sodium pentobarbital to maintain a deep state of anesthesia. A cannula was inserted into the trachea. During the measurements the animals were kept warm with a homeothermic blanket (Harvard). The experiments were conducted in a double-walled soundproof and faradized room (Industrial Acoustics Company).

One of the pinnas was surgically removed and the auditory bulla was exposed and opened. Through the opening in the bulla a silver Teflon insulated wire with ball electrode was inserted and placed at the round window. The electrode lead was glued to the bone and the bulla opening was closed again with ear impression compound (Microsonic). Two silver wire electrodes were placed in the skin: a reference electrode at the nape of the neck and a ground electrode near the contralateral bulla. The AN was exposed through a posterior fossa approach, involving the removal of a small area of cerebellum. The AN measurements were performed with insulated Platinum/Iridium (90/10) ball electrodes mounted on a custom designed and made electrode holder for differential nerve recordings. In a preliminary study different electrode configurations and positions were evaluated. In the actual study, two experiments were conducted with differential ball electrodes placed longitudinally on the nerve, with the active electrode closest to the internal auditory meatus; and two other experiments with a "monopolar" configuration, with the active electrode on the auditory nerve and the reference at the nape of the neck.

Stimulus generation

Stimuli were generated with custom software and a digital sound system (Tucker-Davis Technologies, system 2, sample rate: 125 kHz/channel) consisting of a digital-to-analog converter (PD1), a digitally controlled attenuator (PA5), a headphone driver (HB7) and an electromagnetically shielded acoustic transducer (dynamic electro-acoustic transducer, Radio Shack, 20 Hz - 50 kHz). The transducer was connected with plastic tubing to a custom earpiece coupler. The coupler was fit in the transversely cut ear canal. The acoustic system was calibrated *in situ*, through the custom coupler, with a calibrated probe-microphone (Brüel & Kjær, type 4192, ½-inch condenser microphone and conditioning amplifier Nexus 2690).

Signal recording

The potentials were differentially measured with battery operated low-noise preamplifiers (Signal Recovery, Model 5115 and/or 2x Stanford Research, SR560) located in the shielded room. For the AN, in a preliminary study different electrode configurations and positions were evaluated. In the actual study, the first two experiments were conducted with differential ball electrodes placed longitudinally on the nerve, with the active electrode closest to the internal auditory meatus; and two other experiments were conducted with a 'monopolar' configuration, with the active electrode on the auditory nerve and the reference at the nape of the neck. The RW measurements were performed between the active electrode at the round window and the reference electrode in the nape of the neck. The grounds of the preamplifiers were connected to the electrode at the contralateral mastoid. The signals were filtered (0.03-30 kHz, cut-off slopes 12 dB/oct.) and further amplified with an external amplifier (Dagan, EX4-400) to a total gain of x10 000. All of the relevant signals including the electrical responses, the stimulus waveform, and the synchronization pulses at the start of every probe were visualized on an oscilloscope (LeCroy, WaveSurfer 24Xs); moreover, the responses were sampled with an ADC (TDT, RX8, ~100 kHz/channel, max. SNR 96 dB) and stored on disk for further signal processing. To increase the signal-to-noise ratio (SNR) of the response, the representations (positive polarity & negative polarity) were repeated at least 100 times and thereafter averaged. The responses shown in this study are all averaged.

Stimulus Paradigm

The paradigm was designed with an emphasis on the quantitative extraction of the entangled neurophonic and, importantly, on the verification of the neurophonic. Considerable attention was paid to the cancellation of the RW-CM and to the removal of artifacts. To achieve this goal, two specific properties of the CM and neurophonic were exploded in the

paradigm, the great linearity and non-linearity of the CM and neurophonic, respectively, and the temporary suppression of neural activity in nerve fibers by forward masking.

The paradigm used in this study is illustrated in Fig. 1 by an example of the first halve representation of the stimulus with corresponding averaged electrical responses at the AN and RW. The second halve representation of the stimulus is equal to the first halve but is polarity alternated. One half representation consists of 3 segments. The first segment A (1-151 ms) contains only a probe; the second segment B (151-403 ms) contains a masker followed by a probe; and the last segment C (403-707 ms) contains only a masker. The probes were pure tones with durations of 50, 100 or 150 ms, depending on the experiment. The maskers were pure tones with a frequency equal to the probe frequency (f_P) or a fixed broadband maskers (typically 50-8000 Hz) with a fixed duration of 100 ms. Note that the masker level is the level over the total noise bandwidth. The interval between the masker and the probe was 1 ms and between different stimuli at least 50 ms. Segment C had a larger stimulus-free period, which was required for the adaptation of the probe and was used for the determination of the noise floors in the results. To reduce the spectral splatter at the stimulus transients, the probe and masker were on and off gated with a 1 ms raised-cosine. The probe level (50 or 55 dB SPL) was chosen such that the SNR of the probe and the numbers of averages were adequate and realistic; and the masker level (65 or 70 dB SPL) was chosen such that the probe response was properly masked.

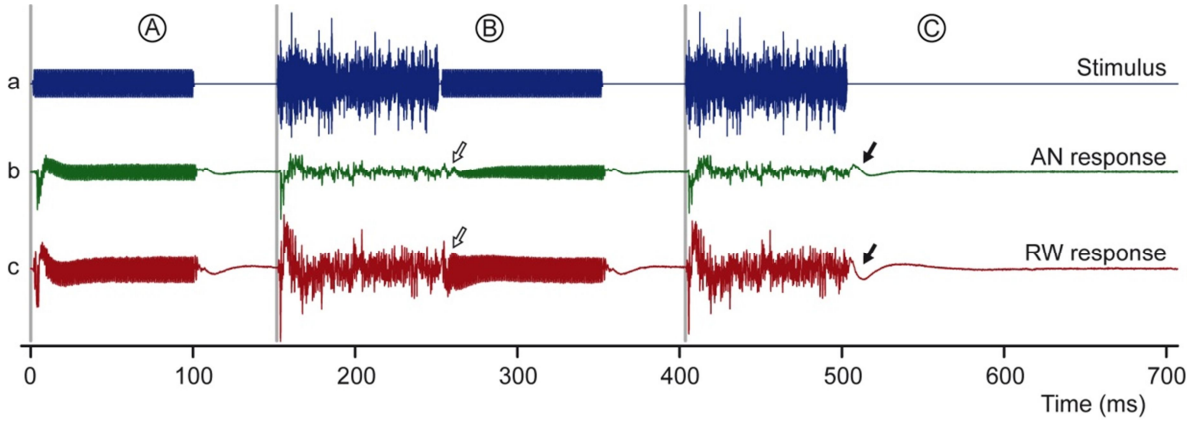


Figure 1. An illustration of the paradigm using the 3 segments of a real stimulus (A) with corresponding averaged responses at the AN (B) and RW (Cs). Section A contains the probe only, B, the probe preceded by a masker, and C the masker only. The probe is a tone of 1 kHz and the masker is a broadband noise. The arrows indicate the offset response of the masker.

Analysis

The responses of the different segments were combined and processed to obtain the required signals. As can be seen in Fig. 1b,c, there is an offset response to the masker (filled arrows 503-540 ms). This component is also present in the response of segment B where it interferes with the response to the probe (open arrows in Fig. 1Bb,Bc). Therefore, the first operation always performed was the removal of the masker's trailing off-set response, by subtracting the response after the masker in segment C (503-653 ms) from the corresponding response in the segment B. In the rest of our analysis, we only make use of these corrected responses, which we refer to as the *masked responses*, and the responses to the probe (segment A in Fig. 1), which we refer to as the *probe responses*.

As mentioned, Fig. 1 shows only half of the stimuli presented: 3 segments identical to that halve but inverted in polarity immediately followed, from which again masked responses and probe responses were obtained. Thus, one stimulus sequence resulted in 2 *pairs* of responses. These are illustrated in Fig. 2Aa (a pair of probe responses to stimuli of opposite polarity) and Fig. 2Ab (a pair of masked responses to stimuli of opposite polarity). From these pairs of responses, we derived a variety of signals and magnitude envelopes which are listed below with their method of calculation: the full significance of these derived signals will become clear in the RESULTS. Except if mentioned otherwise, the amplitudes are presented as a peak voltage.

Within a pair, the two responses were subtracted or added. Subtraction resulted in *stimulus polarity dependent responses* (Fig. 2B). For example, the probe response to positive stimulus polarity (Fig. 2Aa, magenta) was subtracted from the masked response to negative polarity (Fig. 2Aa, cyan) and this difference was halved (Fig. 2Ba). Similarly, a halved difference was obtained for the masked responses (Fig. 2Bb). Summation resulted in *stimulus polarity independent responses*. Thus, Fig. 2Ca shows the sum of the two probe responses, and Fig. 2Cb shows the sum of the two masked responses.

Ultimately, we want to examine the part of the response that can be masked, which is taken to be purely neural. As further explained in RESULTS, this part of the response is obtained by subtracting the masked response (Fig. 2Ab,Bb,Cb) from the probe response (Fig. 2Aa,Ba,Ca), resulting in the *adapted component* (Fig. 2Ac,Bc,Cc). This component contains both phase-locked an non-phase-locked neural contributions: which is particularly noticeable in the polarity-independent onset response, referred to as the compound action potential (CAP), e.g. in Fig. 2Ac,Cc (Goldstein and Kiang, 1958; Kiang et al., 1976; Antoli-Candela and Kiang, 1978). Because we are only interested in the phase-locked neural contribution, we removed the CAP with a non-causal phase preserving high-pass Chebyshev FIR filter

(MATLAB, *filtfilt, fir1*, filter order >200). The cut-off frequency applied to the response pairs of Fig. 2Ac and their difference in Fig. 2Bc was ½f_P and it was f_P for their sum (Fig. 2Cc). In RESULTS, we will argue that the resulting signals (Fig. 2Ad,Bd,Cd) represent a pure neural phase-locked contribution to the neurophonic and refer to it as its decaying part: AN-dNP. Note that the non-decaying part, reflecting sustained phase-locked neural activity, is more difficult to disambiguate from the AN-CM: we return to this contribution in RESULTS (Fig. 3).

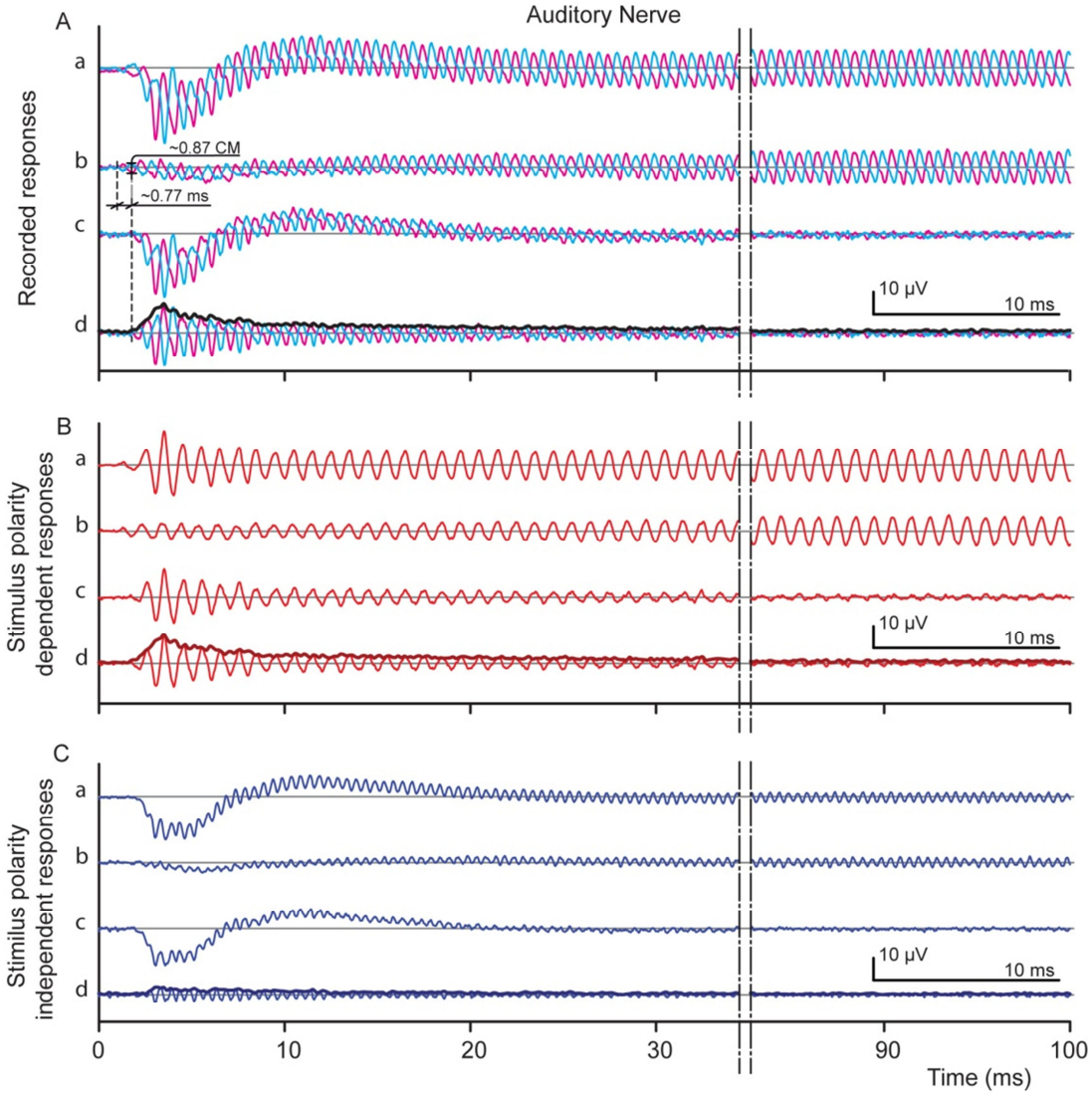


Figure 2. Example of stimulus evoked response averages (n = 128) recorded on the AN. Aa) Pair of raw probe responses for two stimulus polarities. Ab) Same as Aa, but with a preceding tonal masker. Ac) Difference between signal Aa and Ab: the adapted component. Ad) Same as Ac, but with the CAP removed by high-pass filtering. Ba-d) Difference of the responses shown in A. Ca-d) Sum of the responses shown in Aa-d. Probe parameters were: level = 50 dB SPL, frequency = 1 kHz, duration 100 ms. Masker parameters were: level = 70 dB SPL, frequency = 1 kHz. For the final trace in each panel (Ad,Bd,Cd) the envelope is shown as the thick line.

To obtain the time course of the amplitude of the AN-dNP signals, we generally used the Hilbert transformation. The resulting envelopes were de-noised using wavelets (MATLAB, *wden*, settings: soft heuristic SURE thresholding with decomposition at level 4 by symlet8 wavelet), shown as thick lines in Fig. 2Ad,Bd,Cd. For the paired signals (Fig. 2Ad), the envelope was calculated as the amplitude corresponding to the average power of the envelopes of both polarities.

For some signals, we also measured *steady-state amplitudes*: these were measured over the stable and transient-free period of the response (Fig. 2: to the right of the dash-dotted line) using a Fast Fourier Transformation. The steady-state amplitude was determined as the amplitude corresponding to the power calculated as the sum of the spectral powers within ±5% of the probe frequency. For the paired signals (Fig. 2A), the spectrum was taken as the average of both polarities.

For reasons that will become clear in RESULTS, we also obtained amplitudes with a Gabor transform, which is a special case of a short-time Fourier transform (STFT; MATLAB, spectrogram) with a Gaussian window. In this study, the Gaussian window was truncated at $\alpha = 2.5$ and had a fixed defined window length of 6 cycles of the probe frequency. The STFT-window was moved in steps of 100 μsec with a minimum overlap of 75% and a maximum overlap of 98%. The size of the FFT was chosen such that the frequency spacing between the spectral components was 2.5% of the desired frequency. For every time step the center frequency (max. power) was searched within a spectral range of +/-10% around the desired frequency. The magnitude as a function of time was obtained as the amplitude corresponding to the power calculated as the sum of the spectral powers within a range of ± 20% of the desired frequency (= 98.9% spectral coverage of the desired amplitude). In case the total power of multiple harmonics was desired, like for the total harmonic (1st, 2nd, 3rd, and 4th) contribution, then the maximum amplitude was sought of the combined spectral harmonic powers in time. The length and the overlap of the time window and the spectral range for the summation of the total spectral power around the individual harmonic was determined by the lowest harmonic frequency.

Noise floor

For the assessment of the upper frequency limit of phase-locking, the signal noise floor was quantified. It was calculated from the background noise in the artifact free part of segment C (e.g., Fig. 1, 540 - 706 ms). This calculation involved, when possible, the response combinations and operations (e.g. subtraction, STFT, etc.) that were also used to quantify the responses. In case not all operations could be performed, the noise was compensated with a

factor corresponding to these operations assuming uncorrelated noise. For example, the removal of the masker's trailing off-set response (cf. arrows Fig. 1) involves a subtraction which for the noise estimate was replaced by a multiplication with √2. Noise floors for the peak and STFT amplitude were calculated as the 99.75^{th} percentiles (obtained with spline interpolation) of the noise distribution around the mean value of the processed noise plus the mean value itself. The reason for this seemingly insignificant operation was to reduce fluctuations in the baseline due to freaks in the noise.

RESULTS

In this study we focus on the determination of the upper frequency limit of auditory neural phase-locking in population responses. We searched for a method, potentially applicable in humans, that provides a representative measure for the neurophonic. In a first section we illustrate the responses and how we process them to yield a signal dominated by neural activity. Next, a number of candidate measures are examined and discussed. Third, the same method is assessed at the RW. Fourth, the influence of a few stimulus parameters like masker level, probe level, and frequency are investigated. Fifth, the method is verified using TTX. In addition the RW-CM is extracted, verified and the relationship with the neurophonic is discussed. Finally, mean trends are calculated and upper frequency limits and cut-off frequencies are obtained.

The probe frequencies tested varied from 0.3 to 8 kHz and were presented in an arbitrary sequence. We used fixed probe levels of 50 or 55 dB SPL and fixed masker levels of 65 or 70 dB SPL. These levels were determined in a preliminary experiment and chosen such that the SNR was large enough and that the neural masking was sufficiently strong. Except when stated otherwise, the probe and masker levels within an experiment were kept constant. The RW mass-potentials were simultaneous recorded with those at the AN.

Auditory Nerve recordings

A typical result of the recorded probe responses and their processed versions at the AN for a probe frequency of 1 kHz is shown in Fig. 2. The results are divided into three Panels. Panel A contains the raw and some derived signals. Panel B and C contains the transformed (stimulus polarity) dependent and independent signals, with some derived signals similar to those as in panel A.

Raw responses. The paired signals in Fig. 2Aa are the (averaged) responses to probe tones (e.g., segment A, Fig. 1): a pair consists of responses to stimuli with opposite polarity. We indicate responses to the positive polarity in magenta color, and those to negative polarity in cyan. Both responses show two distinct features, a "slow-wave" and a distinct alternating component. The slow wave, here consisting of a negative peak followed by a shallow positive peak, is the compound action potential (CAP); it is the weighted summed extracellular response of synchronized single auditory nerve fibers at the onset of the stimulus (Goldstein and Kiang, 1958; Kiang et al., 1976; Antoli-Candela and Kiang, 1978). The CAP shows a large independence of the stimulus polarity. The other distinct feature, the alternating component, shows strong phase coupling with the underlying polarity alternated stimuli. The tight phase coherence between these responses is often exploited to separate the (stimulus polarity) independent part of the response, like the CAP, from the dependent part.

The polarity independent response is represented by Fig. 2Ca and is calculated as half the sum of the corresponding paired signals in Fig. 2Aa. Apart from the CAP, an alternating component is present at double the frequency of the probe. It is believed that this component is dominated by the second harmonic of the AN-NP. This second harmonic is caused by the rectifying effect of synaptic transmission and is therefore not explicitly present in the CM. The reason why the second harmonic occurs in the polarity independent response is because the even harmonics in the polarity alternated phase-locked responses are identical and are summed, while the odd harmonics are cancelled. The opposite is also true, the odd phase-locked harmonics are retained in the polarity dependent response (obtained by subtraction of the paired responses in Fig. 2Aa) and the even harmonics cancelled.

The polarity dependent response is shown in Fig. 2Ba and is clearly dominated by an alternating component at the stimulus frequency (1st harmonic); the CAP and even harmonics are not visibly present (this was spectrally verified, not shown here). In contrast to the polarity independent response, the dependent response is likely not entirely of neural origin: some stray CM might be present like at the RW. To verify this, a strong forward masker was applied.

Masked responses. The paired signals in Fig. 2Ab are the responses to probe tones preceded by a strong tonal masker. The corresponding polarity dependent and independent masked responses were calculated and are shown in Fig. 2Bb,Cb. The presence of a masker has a clear depressant effect on the probe response: both the CAP (Fig. 2Ab and Cb) and the alternating component (Fig. 2Ab,Bb,Cb) are reduced but the reduction is temporary and the signals recover to their non-masked amplitude toward the end of the responses (Fig. 2, 90-100 ms). Comparison of probe and masked responses show that at least 85% of the CAP

(Fig. 2C) and 75% of the alternating component (Fig. 2B) are masked. The observation that forward masking is only partial, i.e. that there is still a remaining signal even at the beginning of the response, may indicate that not all neural activity is suppressed by the preceding masker, or that the response contains a non-neural non-adapting component.

Adapted component. The net result of masking is shown by subtracting the masked response from the probe response. For example, subtraction of Fig. 2Bb from Fig. 2Ba results in Fig. 2Bc. The small amplitude in the later part (vertical dashed-dotted lines) of these subtracted signals indicates the recovery of the response. Since we are only interested in the phase-locked components, a high-pass FIR-filter was applied to remove the CAP and other not relevant components without introducing other side effects like ringing and phase distortions (METHODS). The cut-off frequency was positioned at halve the frequency of the harmonic considered. The signals after filtering are represented at the bottom of each panel (Fig. 2Ad,Bd,Cd); the time course of their amplitudes is indicated by the thick line. These three envelopes are the signals of interest in most of the sections that follow.

The course of these envelopes depicts three distinct regions: a region near the onset where the envelope has a steep rising slope, followed by a region with a steep declining slope, which thereafter merges into the last region with a less steep slope. Qualitatively, the envelopes have the typical appearance of a high-pass filtered (DC-free) post stimulus time histogram of primary auditory nerve fibers, but with an additional exponential decay as a result of the forward masking. This similarity in shape, as well as the full recovery, suggests that the adapted component is at least dominated by neural contributions.

Strong evidence for the presence of two response components which differ in adaptation and delay is given by the time lag of the adapted component with respect to the probe and masked responses. This is illustrated (dashed lines) in Fig. 2A, which shows a minimum estimated lag of 0.77 ms (Fig. 6 also illustrates this point at a finer timescale). The size of this lag is roughly compatible with a synaptic delay combined with some additional conduction delay. The most obvious explanation for the time lag and the observation that masking is only partial and not complete, is that the responses are contaminated by stray current from the cochlear receptors. Taken together, these results provide strong evidence that the maskable response only consists of phase-locked neural contributions; henceforth we refer to that component as the decaying AN-NP (AN-dNP).

The presence of stray CM (AN-CM) is not entirely unexpected. The AN in cat is relative short and is embedded in conductive surroundings so that it is extremely difficult, if possible at all, to completely electrically isolate the AN from these surroundings. The time lag in the neural response provides a window to quantify the AN-CM. Applied to the 0.77 ms

window in the example of Fig. 2, the estimated AN-CM amplitude is ~2.6 µV. We verify this method further below in the section "TTX".

Quantification of the neurophonic

Our main objective is to find a representative measure for the AN-NP that enables determination of an upper frequency limit. The amplitude of the 2nd harmonic of the AN-NP, which can be obtained from the steady-state part of the polarity independent response (e.g., Fig. 2Ca) as described in the previous section ("raw responses"), has been interpreted as being indicative of phase-locking (Snyder and Schreiner; Lichtenhan et al., 2013). It is easily obtainable and contains almost no CM, but, for reasons that we return to below (RESULTS, Fig. 13; DISCUSSION), it is unsuitable to assess the upper frequency limit. To obtain a better metric, we first show that the steady-state and adapting response components can be related to each other in a manner that is coherent and consistent with known neural properties.

Steady-state estimation. We first compare the unmasked response and the AN-CM (estimated as discussed in the preceding section). This is graphically illustrated in Fig. 3.

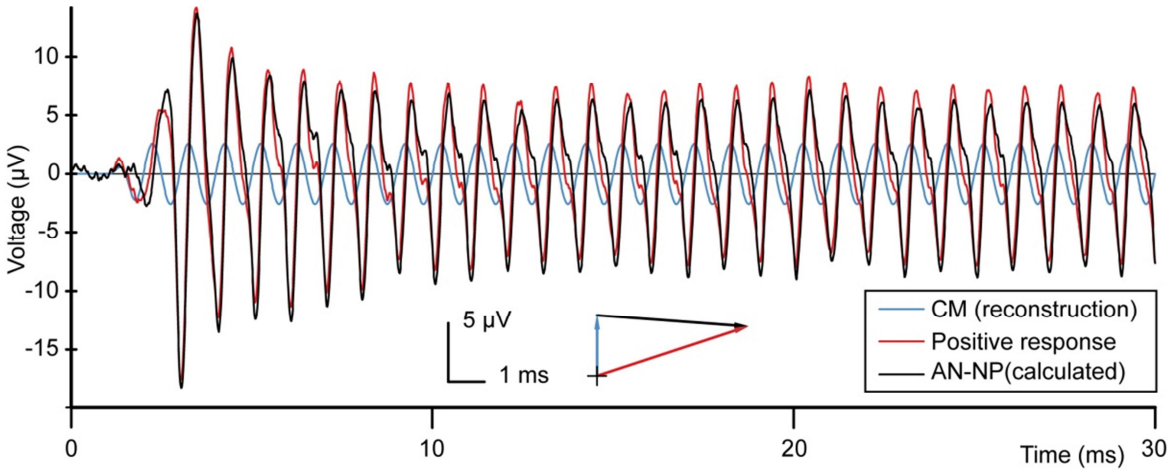


Figure 3. Estimation of the AN-NP, by subtracting the CAP-filtered positive probe response with the reconstructed AN-CM. The vector diagram represents the vector subtraction of the steady-state amplitudes between the positive probe response and the estimated AN-CM.

The red trace shows one of the raw probe responses (the magenta trace in Fig. 2Aa, but filtered to remove the CAP), in the time domain and as a vector. The blue signal is the reconstructed AN-CM based on the phase and amplitude before and at neural onset (Fig. 2Ab,c, dashed lines) taking into account signal gating. The AN-NP is here determined as the difference between the probe response and the estimated AN-CM and is indicated as the black signal. The AN-NP amplitude is then measured from the steady-state region. An equivalent

method is a vector subtraction, which gives an AN-NP of [7.3 µVp, ∠-73°] – [2.6 µVp, ∠0°] = [7.0 µVp, ∠-94°]. Thus, the AN-CM is almost a third in amplitude and at right angle re. the AN-NP. Note that this -94° phase difference is consistent with the time delay noted above, when combined with a phase reversal between AN-CM and AN-NP. Indeed, a phase reversal is noted when comparing the onset response waveforms of the two components, e.g. a zoom-in (not shown) shows that Fig. 2Bb starts positive going while 2Bc starts negative going. Thus the phase predicted for the AN-NP equals -94° - 180° = -274°, which is close to the phase delay predicted by the onset lag (0.77 ms x 360 ° / 1 kHz = -277°). The good agreement indicates that the response measured on the AN can be reasonably described as the vectorial sum of an AN-CM and AN-NP with a fairly consistent phase relationship at the onset and steady-state.

Adapted component. It was argued in the previous section that the decaying neurophonic (AN-dNP) reflects a purely neural contribution (Fig. 2Ad,Bd). If this is so, we should be able to relate that contribution to the steady-state estimate just derived. The red trace in Fig. 4 shows the positive response of Fig. 2Ad (magenta). This is the AN-dNP i.e. the adapting neural contribution. It is overlaid on the AN-NP estimated with the steady-stated method of Fig. 3 (black trace in Fig. 3 and Fig. 4). At response onset, these curves are in good agreement. They have the same phase and a similar intrinsic adaptation at the onset (before 10 ms), but the AN-dNP reflects an additional exponential decay due to the recovery of masking.

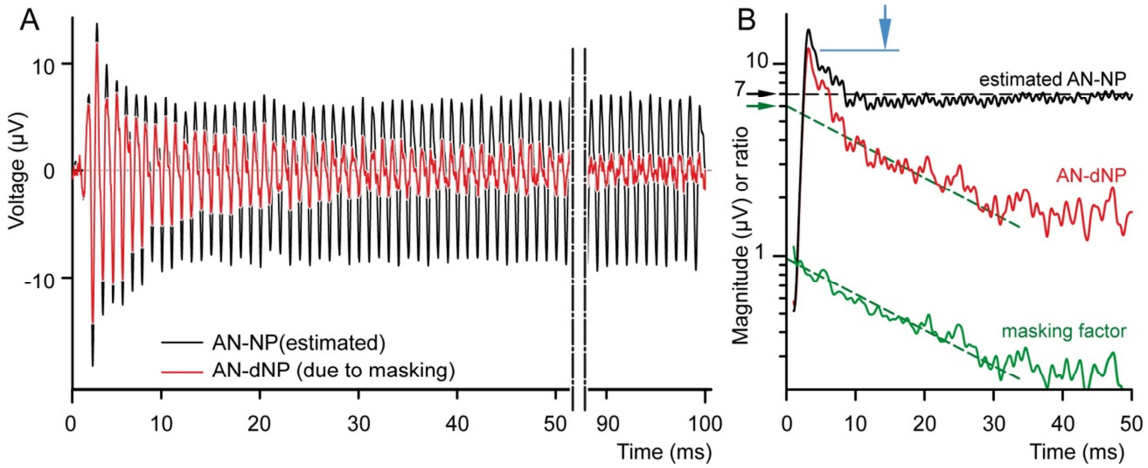


Figure 4. Comparison between the decaying AN-NP (AN-dNP, red) and the estimated AN-NP from Fig. 3 (black). Only the positive polarity is considered. A) Overlay in the time domain. B) Magnitudes of the signals in A together with the masking factor of the AN-dNP induced by the exponential recovery of masking (lower green line, calculated as the ratio of the magnitudes and slightly smoothed with a non-causal moving average filter). The blue arrow is the peak-magnitude of the AN-dNP; the black arrow is the estimate of the steady-state amplitude (7 µVp); and the green arrow is the extrapolated amplitude based on the exponential recovery of masking.

Fig. 4B shows the instantaneous amplitudes of the waveforms shown in Fig. 4A on a log-lin scale. The ratio of these two amplitudes is the masking factor (green curve): a factor of 1 indicates 100% masking while zero would indicate full recovery. The curve shows an exponential decay (linear trajectory fit by dashed line) up to ~30 ms with a time constant (slope) of 18 ms; thereafter the time constant increases to a much higher value (not shown), consistent with the literature ((Chimento and Schreiner, 1990), mean time constant for the rapid recovery: 16.2 ms and for the slow recovery: 125 ms for a tone and masker of 800 Hz). Note that application of this decay to the AN-dNP (green dashed line through red trace, Fig. 4B) provides an extrapolated value at 0 ms which is close to the steady-state value estimated previously (Fig. 3: 7.0 μVp). This agreement is not trivial, since the waveforms of Fig. 4A were derived from different response components.

The extrapolated value at 0 ms could be used as a metric of neural phase-locking. However, it would only be available for responses with a clear slope of recovery, and not for low SNR signals where only the intrinsic adaptation is significantly present. Because of our interest in determining the upper limit of neural phase-locking, where the SNR is intrinsically poor, another more robust measure is needed. We therefore selected the peak-magnitude of the AN-dNP (vertical arrow in Fig. 4b). This measure is robust because the peak is easily detectable and is inherently the largest instantaneous signal. A small disadvantage is that it does overestimate the steady-state value of neural phase-locking.

Round Window Recordings

In the previous section we assessed the responses at the AN for a probe frequency of 1 kHz and found a suitable measure of the neurophonic. As mentioned in the introduction, we seek a method which is also applicable in humans, where recording from the AN is not possible except under neurosurgical conditions, but where recording from sites near the RW is possible. Therefore, we examined whether the method would be applicable to responses recorded at the RW.

Fig. 5 shows the responses at the RW recorded simultaneously to the AN responses in Fig. 2 and arranged in the same format. The probe responses (Fig. 5Aa) and their difference (polarity dependent responses, Fig. 5Ba) and sum (polarity independent responses, Fig. 5Ca) show the same basic features as for the AN. Again, an alternating component and a CAP are present, but there are some differences with the AN responses. First, the CAP (Fig. 5Ca) is larger and shorter. Secondly, the alternating component has a much larger amplitude (Fig. 5Aa, different ordinate scale than Fig. 2A) than at the AN and has a different onset with a local depression around 5 ms, which is more pronounced in the dominant 1st harmonic

component in panel B. Third, the forward-masked responses (Fig. 5Ab) differ strikingly from the AN responses (Fig. 2Ab): there is no decaying slope and the initial response is actually larger than in the unmasked condition and decreases with time. Masking caused a decrease in CAP magnitude and second harmonic amplitude (Fig. 5Cb) similar to the AN. The apparent opposite effects of forward masking on RW vs. AN and the local depression around 5 ms in the RW probe response are due to the interplay of a large non-maskable RW-CM and the auditory round-window neurophonic (RW-NP).

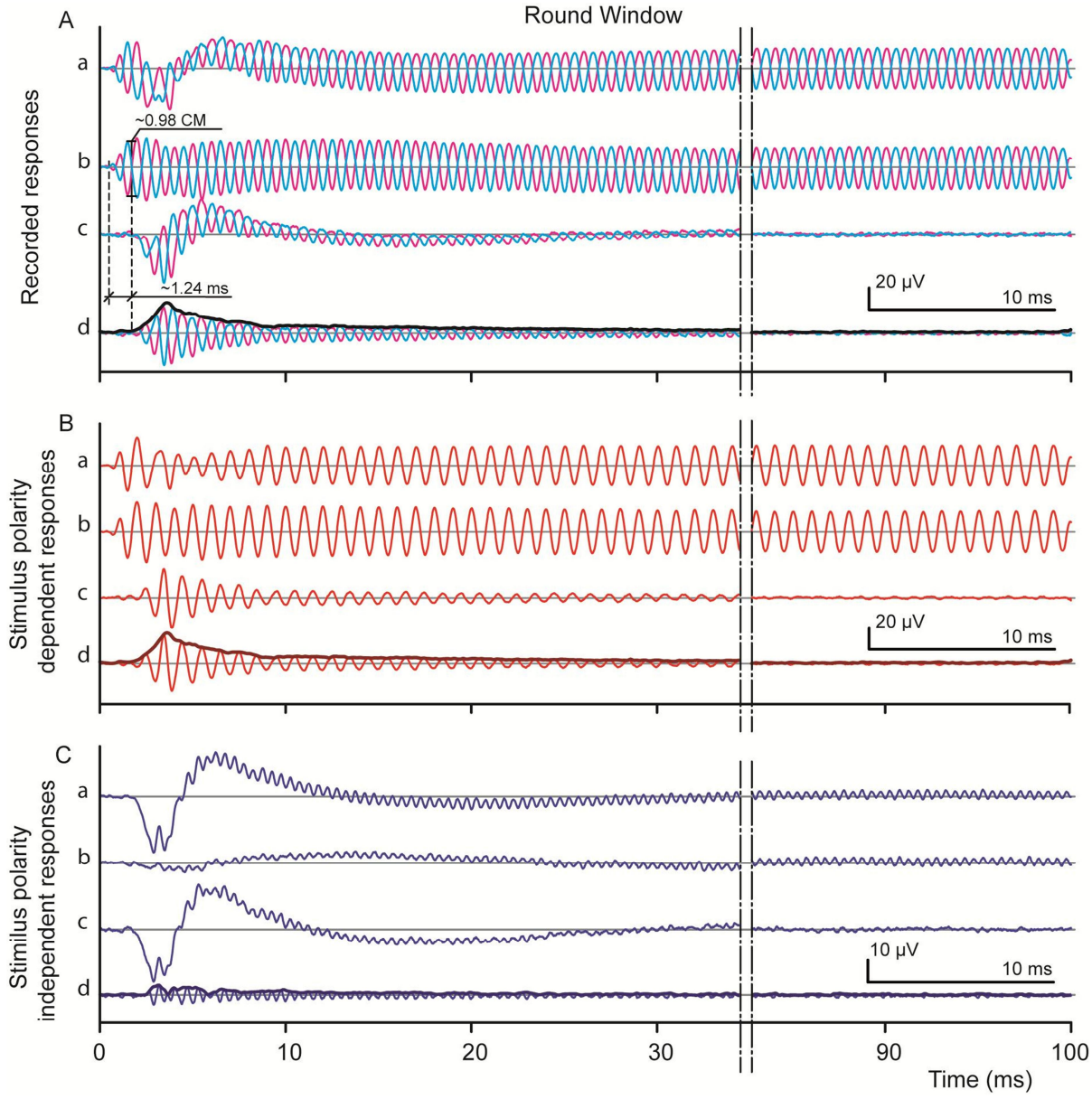


Figure 5. Similar to Fig. 2, but for simultaneous recordings at the RW.

The adapted component (Fig. 5Ac,Bc,Cc) and its CAP-filtered version (Fig. 5Ad,Bd,Cd) reveal the true extent of the RW-NP: despite its wildly differing probe and masked responses, the RW has an adapted component which is surprisingly similar to the AN-dNP. We refer to this component as the round window decaying neurophonic (RW-dNP).

Time lag difference. Similar to the AN, there is a noticeable onset lag (1.24 ms) between the probe and masked responses (Fig. 5Aa,b) and the adapted component (Fig. 5Ac). This delay is considerably larger than for the AN (0.77 ms, Fig. 2Ab) and made us examine the onset regions of these responses more closely. Fig. 6 shows overlays, freely scaled in amplitude, of the RW and AN responses. The traces in Fig. 6A show the polarity dependent masked responses of the AN (green, as in Fig. 2Bb) and RW (red, as in Fig. 5Bb): these traces show the onset of the receptor potential contribution (AN-CM and RW-CM). The traces in Fig. 6B show the decaying neural components i.e. the AN-dNP (green, as in Fig. 2Bd) and the RW-dNP (red, as in Fig. 5Bd). To our surprise, the phase differences are larger for the CM traces than for the neural traces. The latter show a time lag of ~100 μs; the mean neural onset is indicated by the vertical dashed dotted line (Fig. 6B). The onset of the 2nd harmonic and other polarity independent components (Fig. 6C) is similarly delayed. We found that these phase and time differences between AN and RW could differ depending on probe frequency, SPL, and animal, but a full treatment of these differences is beyond the scope of the present study. The main point here is that, similar to the AN recordings, the responses recorded at the RW display two components which differ in onset delay and adaptation and which are consistent with generators in receptors and nerve fibers.

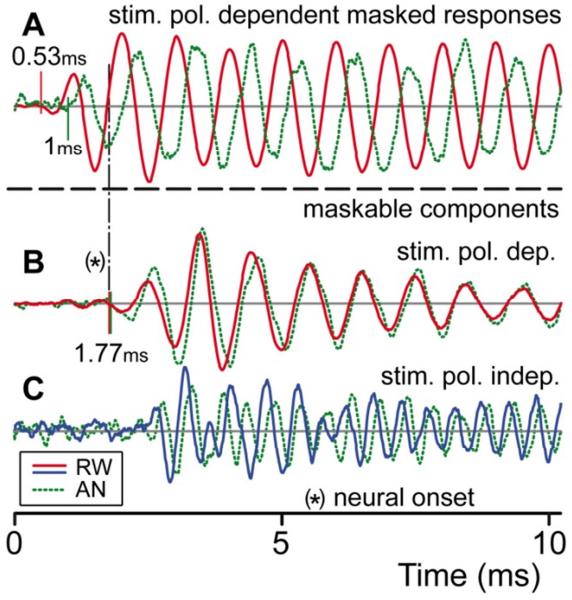


Figure 6. Comparison of the onset regions of responses recorded on the AN (dotted lines) and RW (solid lines). (a) Shows the stimulus polarity dependent masked responses, which at onset reflect CM. Panels B and C show adapted, i.e. neural, components: polarity dependent (b) and independent (c). The corresponding onset values are indicated by the vertical lines. Note that, for reasons of comparison, the amplitudes are individually and arbitrarily scaled; only the polarity independent signals have the same scale.

Dependencies

In the previous sections we explained and illustrated a method to quantify neural phase-locking measured at the AN and RW. In this section we use the method to investigate the dependence of the decaying neurophonic on level and frequency. The main goal is to determine the decrease and upper detectable frequency limit of the auditory neurophonic at the RW, since this type of recording can be applied to humans, and also at the AN, as verification.

Masker Level dependence. Before examining frequency effects, it is important to evaluate the parameters resulting in optimal masking. We investigated the influence of masker level on the peak-magnitude of the AN-dNP. Fig. 7 shows the influence of 6 masker levels on the time-course of the magnitude of the masked responses (panel A and C) and corresponding AN-dNP (panel B) and RW-dNP (panel D). The traces shown are the magnitudes as previously illustrated with the solid black traces in Fig. 2Ad and Fig. 5Ad. The probe level and frequency for this example were 55 dB SPL and 1 kHz, respectively. To visually emphasize masking, the time scales of the panel are compressed by using a logarithmic time scale. Besides the magnitudes of masked AN responses for different masker levels, Fig. 7A also shows an unmasked response (thick blue curve) with its typical intrinsic adaptive shape (the other unmasked responses, measured for each masker level, Fig. 1, are not shown but are similar in shape). Increase of masker level increasingly reduces the response, but the reduction is maximal at about 10 dB above the probe level (i.e. 65 dB SPL for this example). Note the early unmaskable component, indicated by the vertical dashed line at neural onset. This is consistent with our earlier argument (section "adapted component") that this component consists largely of AN-CM. The AN-dNP magnitudes corresponding to the traces of Fig. 7A, i.e. the vectorial differences (cf. Fig. 3) between the unmasked and increasingly masked responses in panel A, are shown in Fig. 7B. The amplitude of the AN-dNP increases proportional to masker level and reaches a maximum (here 20 μV at ~55 dB SPL, arrow).

Fig. 7C show the corresponding traces for the RW responses simultaneously recorded with those in A. Again there is an unmaskable early component (preceding the vertical dashed line), which is much larger than in Fig. 7A. In accordance with the analysis in the preceding section, we label this component as the RW-CM and interpret the subsequent capricious time course of the responses as a reflection of the interaction of the different phases and amplitudes of the RW-CM and RW-NP, which gradually smoothes out with increasing levels of masking. Note that the effect of increasing masking on amplitude is mostly in a direction opposite to that in Fig. 7A: the RW masked response actually increase with increase in masker level. Fig.

7D shows the RW-dNP traces. In contrast to Fig. 7C, these traces are not capricious and are strikingly similar to those of the AN-dNP (Fig. 7B), but systematically 2.3 times larger.

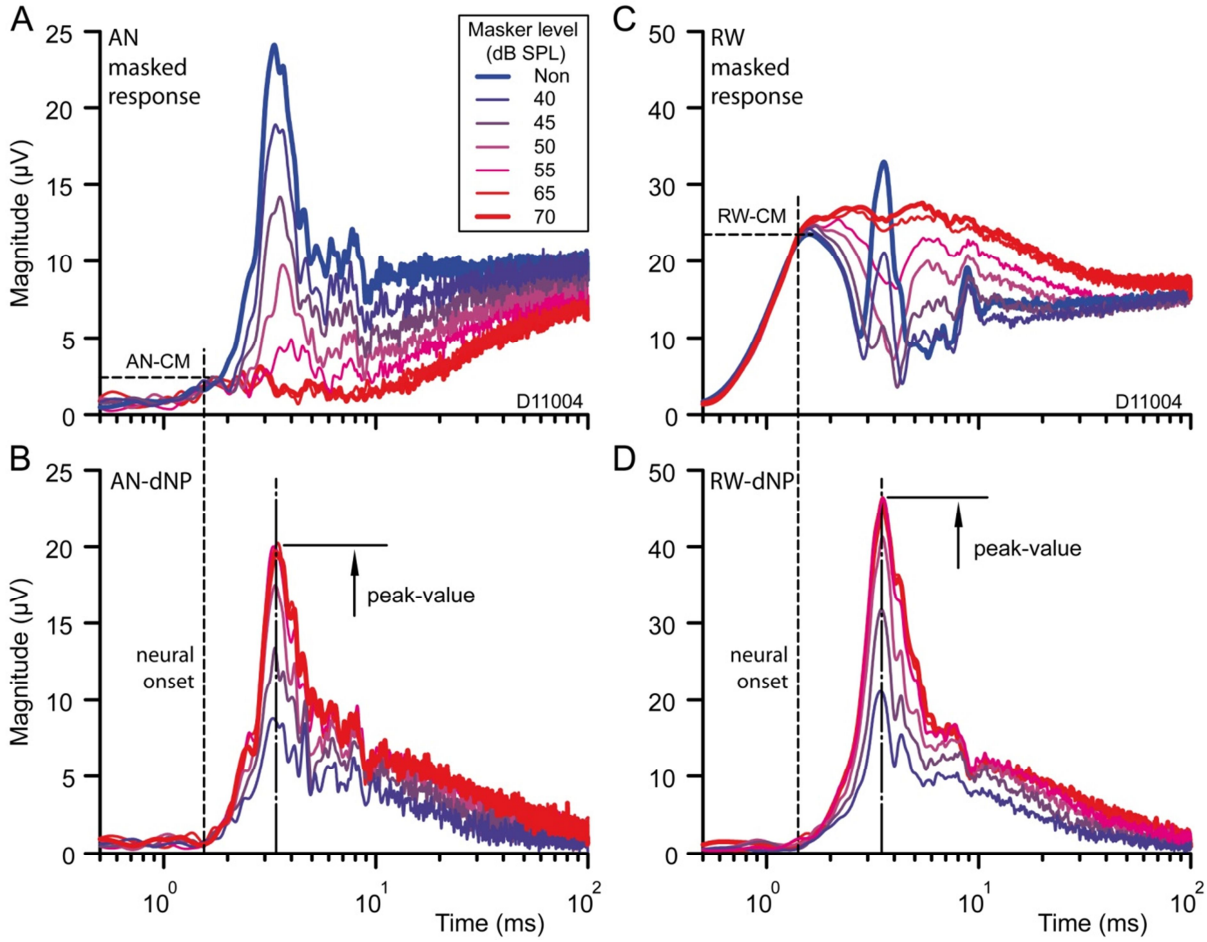


Figure 7. Dependence of the response magnitude time-course on masker level. A,C: masked responses of the AN and RW. One unmasked response is shown with the thick blue line. B,D: decaying auditory neurophonic calculated from the responses in A,C as the vectorial difference between masked and unmasked responses at a given masker level. The estimated AN-CM and RW-CM at neural onset are indicated in panels A and C by the dashed lines. The peak-magnitude of the decaying auditory neurophonic are indicated by the arrows. The probe frequency and level are fixed and equal 1 kHz and 55 dB SPL; the masker was a broadband noise (50 Hz - 8 kHz). Note that this experiment is a different experiment (animal) than those of Fig. 2 and Fig. 5.

To summarize, the example of Fig. 7 illustrates several important features. It shows the tight relation between the neurophonic measured at the RW and the AN, and the ability of the method to reject the CM, independent of masker level. It also shows that masker levels of about 65 dB are sufficient to make optimal use of forward masking towards characterization of the neurophonic. Not illustrated, we also examined the effect of probe level. These results

again show a marked difference in behavior between CM and dNP, which will be illustrated in the final section of RESULTS, but which are not important for the presentation of the results on frequency dependence, which are described next.

Frequency dependence. Fig. 8 presents the time course of the decaying auditory neurophonic for a wide range of probe frequencies at the AN (panel A) and RW (panel B).

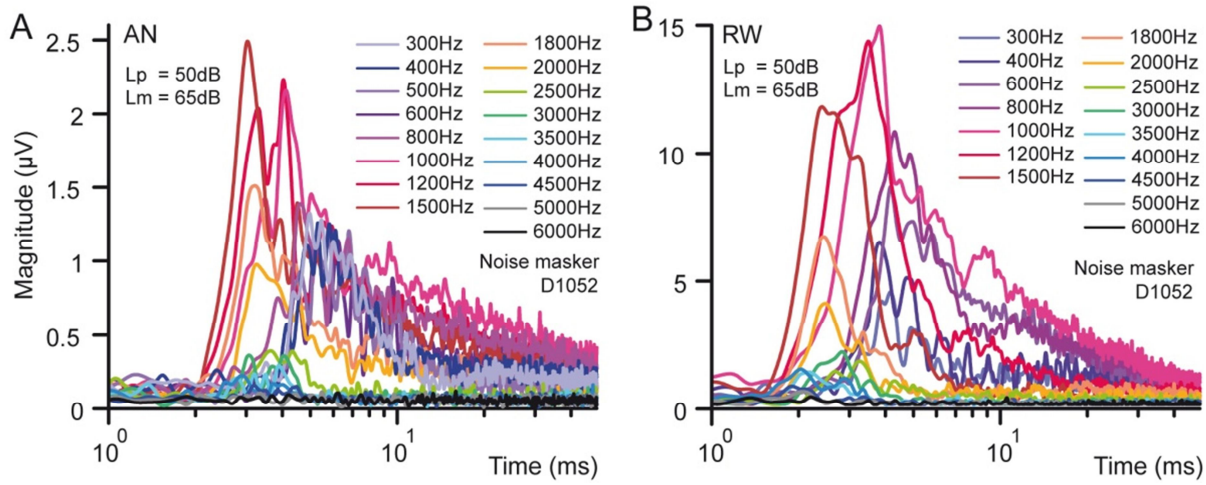


Figure 8. Effect of probe frequency on simultaneously recorded AN-dNP (A) and RW-dNP (B). The probe tone level was 50 dB SPL and the masker was a broadband noise fixed to 65 dB SPL.

Regardless of frequency and measurement location, the time course of the magnitudes is always similar, showing an early peak followed by an exponential decay. For frequencies with low SNR, the tail was not distinguishable from the background noise; in some cases only the region around the peak remained. Two distinct features were frequency dependent: the (peak-) magnitude and the delay of the rising slope. The delay of the rising slope decreased with increasing probe frequency. This trend is not further quantified here but is consistent with cochlear physiology i.e. shorter latencies for high than for low frequencies, as also evident in CAP responses (Eggermont, 1976b), single-fiber responses (Kiang et al., 1965), and cochlear mechanical measurements (Eggermont, 1976b; Robles and Ruggero, 2001).

Regarding the peak-magnitudes, these are depicted in Fig. 9 as a function of probe frequency.The trend lines of the AN (panel A) and RW (panel B) show a similar course: het magnitude increases with frequency until it reaches an absolute maximum, thereafter it decreases sharply until it is bounded by the background noise. The baseline for the measurement is indicated by the pimped dashed line and is the noise floor derived from the stimulus response-free part at the end of the stimulus sequence (Fig. 1C, and see METHODS). The detectable upper frequency limit is here chosen as the location where the

curves merge; for the example shown the limit is 5 kHz for the AN and is undefined for the RW. A higher frequency limit at the RW than at the AN was not always found, though a small trend is observed (see below).

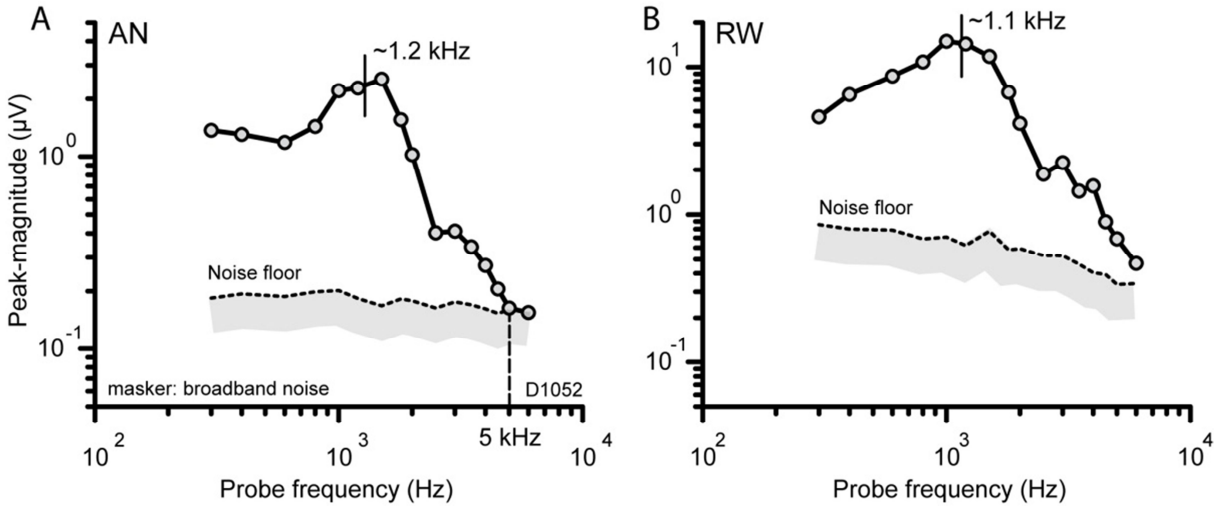


Figure 9. Peak-magnitudes from the AN-dNP (A) and RW-dNP (B) in Fig. 8 as a function of probe frequency. The baseline is indicated by the dashed-shaded line and is the noise floor in the measurement. The estimated centers of the blunt peaks are indicated by the vertical lines and are associated with the cut-off frequencies of the low-pass slope. The (assumed) upper frequency limits are indicated by the dashed vertical lines, and are obtained at the intersection with the noise floor. The probe level was 55 dB SPL; the masker was a broadband noise masker (50-20000 Hz) with a spectral level of 65 dB SPL.

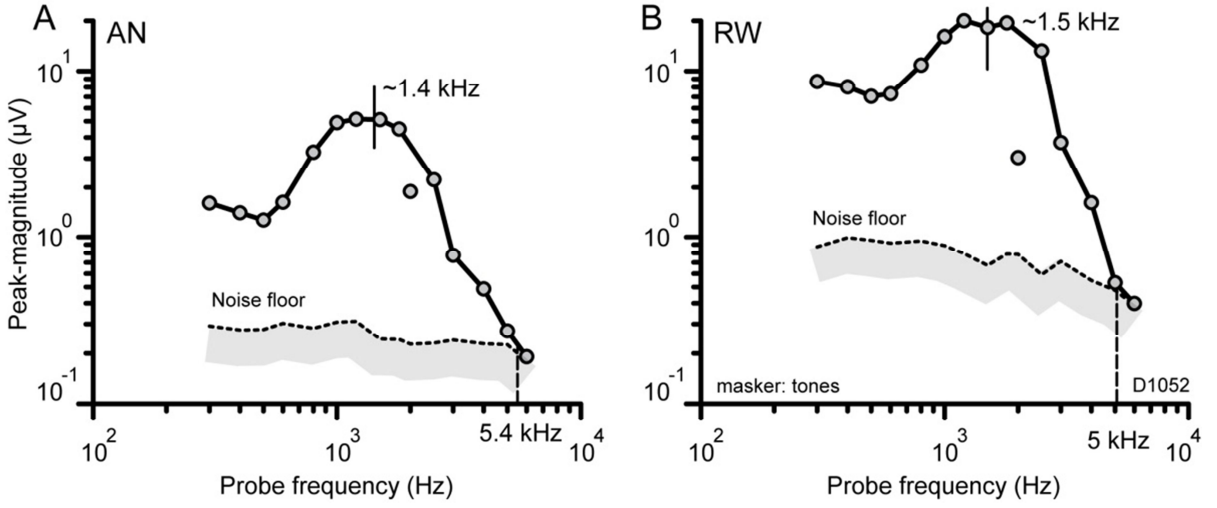


Figure 10. Similar to Fig. 9, but for a tonal masker of 65 dB SPL. Note that this measurement session was conducted in the same animal but at a different time. The disconnected dot is considered to be an outlier.

Another series of measurements in the same animal but for a tonal masker is shown in Fig. 10. The peak-magnitudes show a trend similar to Fig. 9, but for this session the limit was 5.4 kHz for the AN and 5.0 kHz for the RW. We visually mark a center frequency for the broad peak of the general bandpass characteristic in these figures. For the two measurement sessions of Fig. 9 and Fig. 10, these center frequencies were similar for the AN and RW and were between 1.1 and 1.5 kHz.

In broad terms, the characteristics of Fig. 9 and Fig. 10 show a similarity to limits of phase-locking in the auditory nerve (Rose et al., 1967; Johnson, 1980). This observation, combined with the fact that the general shape of these characteristics is very similar for the two recording locations, holds the promise that it may be possible to estimate neural phase-locking from mass potential recordings at the middle ear, including human. But there are several issues. First, our method is based on tests at a probe tone of 1 kHz (Fig. 2-7): we assume that our extraction procedure is also valid at high frequencies near the upper limit, where the signals are inherently too small to allow thorough testing. We further test this assumption below (Fig. 13). The second issue is that overviews of neural phase-locking in single fibers are generally generated by plotting maximal vector strength at a fiber's characteristic frequency (CF), while here we measure the contribution of neurons of all CF to a probe fixed in frequency and level. Differences in phase synchronization across fibers; changes in vector strength magnitude and phase with frequency relative to CF; spectral content of the contributions of single action potentials to the mass potential; etc. are expected to affect the overall shape of the transfer characteristics as measured in Fig. 9 and Fig. 10. Detailed cochlear and neural modeling (cf. Antoli-Candela and Kiang, 1978; Wang, 1979) is required to take such factors into account, which is beyond the goal of the present paper. Despite these factors, congruence between the upper limit of neural phase-locking in the mass potentials and single fibers is expected – within the limits imposed by the noise floor. The shallow high-pass behavior in Fig. 9 and Fig. 10 is not present in population plots of maximal vector strength in the cat auditory nerve (Johnson, 1980; Rhode and Smith, 1985; Joris et al., 1994a) and must attributable to such factors. The third issue regards the fact that the results in Fig. 9 and Fig. 10 are obtained with the total maskable response (e.g. Fig. 2Ac and Fig. 5Ac), so including all maskable harmonics. This is in contrast to typical measures for phase-locking in AN fibers, like the vector strength, which is a kind of synchronization index based on the 1st harmonic in the neural response, but also to studies of mass potentials which make only use of its 2nd harmonic to differentiate neural from hair cell contributions (Snyder and Schreiner, 1984; Lichtenhan et al., 2013). We therefore now examine the frequency dependence of different harmonic components.

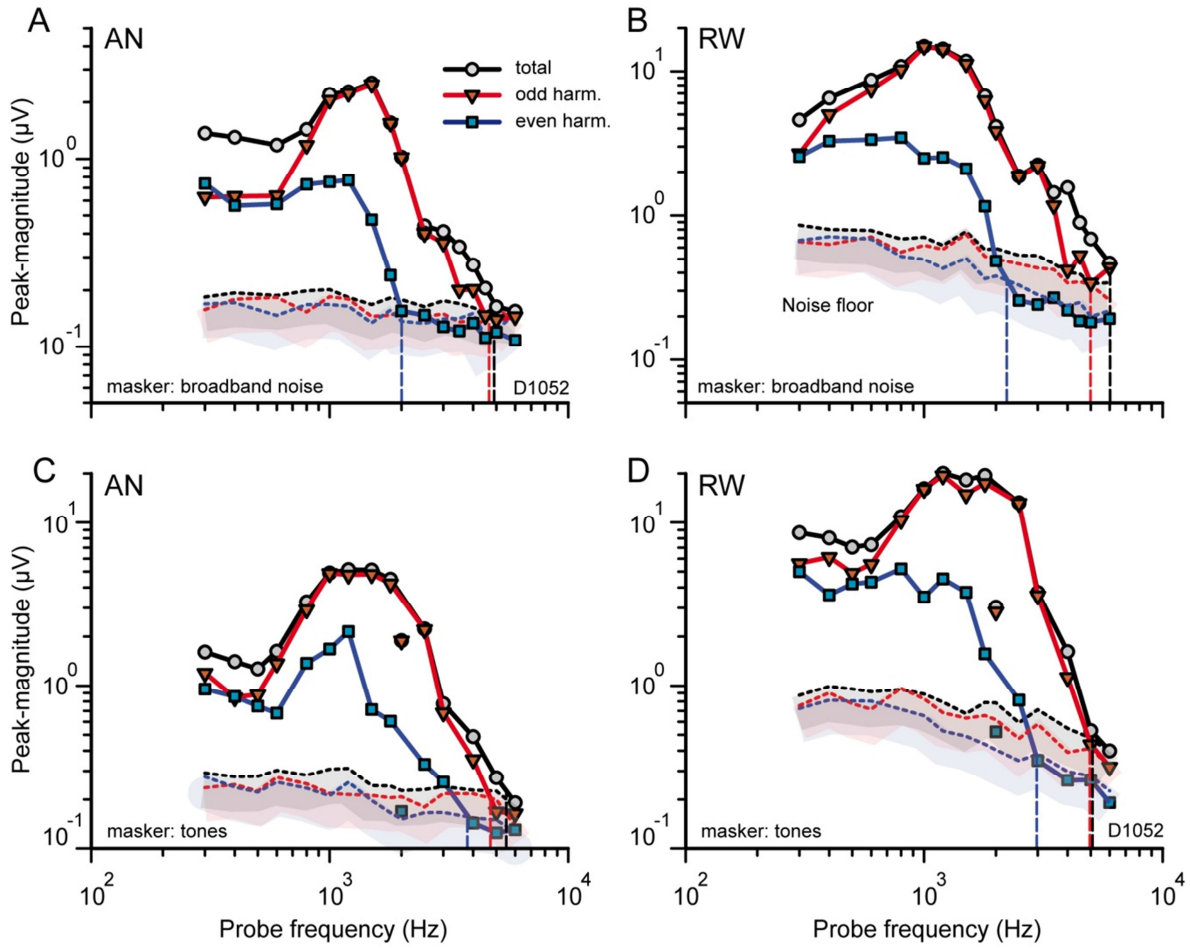


Figure 11. Normal and decomposed AN-dNP and RW-dNP peak-magnitudes of Fig. 9 and Fig. 10. The peak-amplitudes of the stimulus polarity dependent and independent adapted components, labeled according to their dominant harmonics. The noise floors of the measurements are indicated by the corresponding dashed and shaded lines. The disconnected dot is considered to be an outlier.

Harmonic separation. We extracted the peak-magnitudes of the maskable odd (dominated by the 1st) harmonics and even (dominated by 2nd) harmonics using the decomposition of the subtracted i.e. stimulus polarity dependent responses (e.g. Fig. 2Bd and Fig. 5Bd) and of the summed i.e. stimulus polarity independent responses (e.g. Fig. 2Cd and Fig. 5Cd). Fig. 11 again shows the AN-dNP and RW-dNP peak-magnitudes also shown in Fig. 9 and Fig. 10. In addition, the odd and even harmonic responses are represented by the red triangles and blue squares, respectively. Below 600 Hz, the amplitudes of the odd and even harmonic responses are comparable. Above 600 Hz, the total neurophonic is dominated by the odd harmonic response; the even harmonic response has a similar course but is up to 20 dB lower. As a consequence, the maximum detectable frequency limit of the even harmonic response is well below that of the odd harmonic response, which in its turn has a slightly lower detectable limit than the total response. Similar trends were found in other

experiments. More examples are shown in Fig. 12: the frequency limits of the experiment represented by panels A and B are quite similar to those in Fig. 11A,B. The limits for the experiments in Fig. 12C,D are lower than the others.

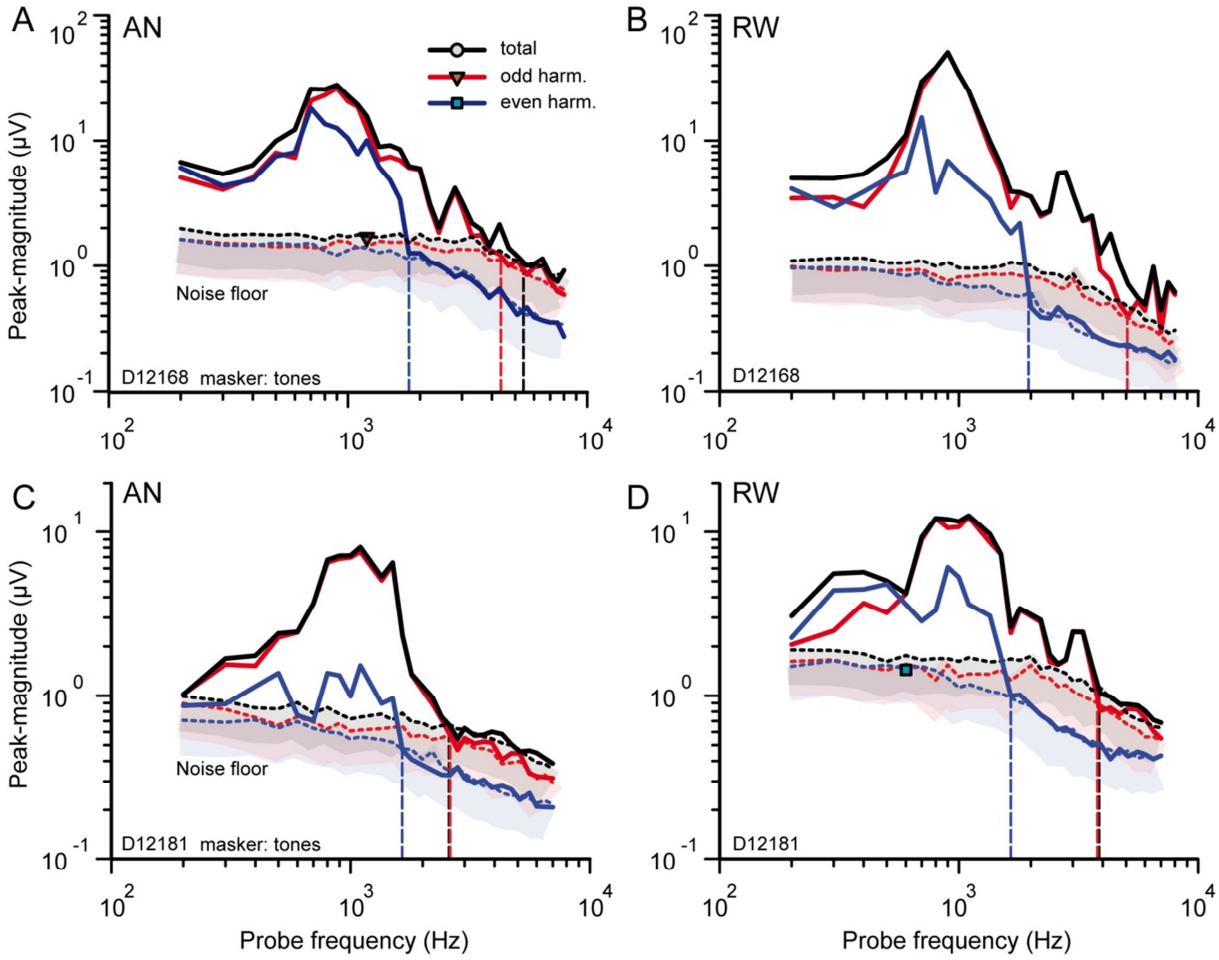


Figure 12. Same as Fig. 11, but for two other experiments. The AN recordings were in both experiments measured in 'monopolar' configuration, between the AN and the nape of the neck.

In order to address the first issue mentioned above, we compared the peak-magnitude of the even harmonic response with its steady-state amplitude. The reasoning is that we have two measurements at our disposal to estimate a purely neural contribution, which we can compare with each other. On the one hand there is the peak magnitude AN-dNP obtained via forward masking, which we have used so far and which we validated at 1 kHz; on the other hand there is the second (and fourth etc.) harmonic which is generally thought to be a purely neural component. This component can be measured in the steady-state response and, if indeed purely neural, provides a means to measure neural phase-locking without requiring

forward masking and. Fig. 13 shows the 2nd and 4th steady-state harmonics as the green and brown curves; the corresponding noise floors are indicated by the dashed lines. The blue curves are the peak-magnitude of the even harmonic response, as also shown in Fig. 11 (blue curves).

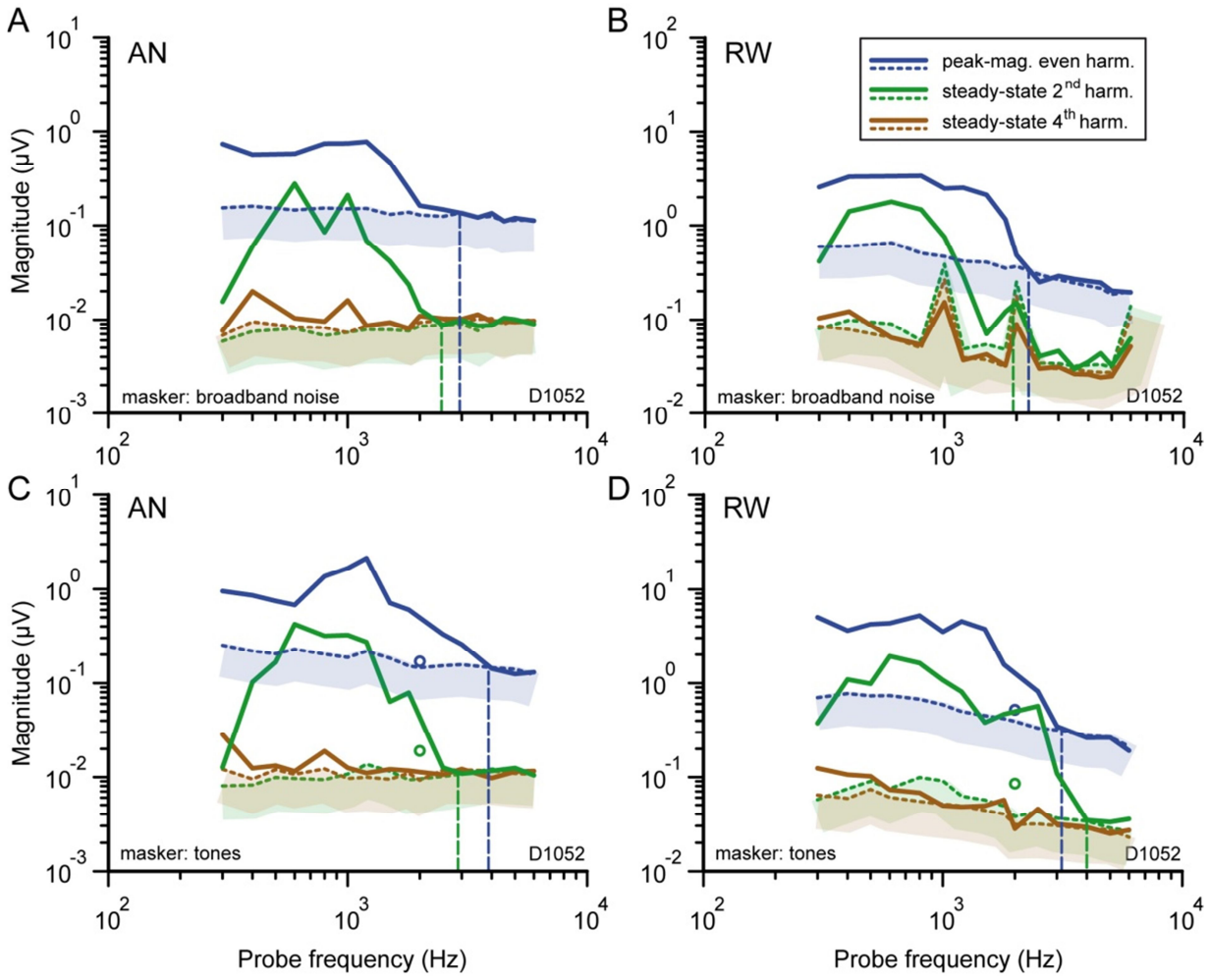


Figure 13. Comparison between the 2nd harmonic peak-magnitude and steady-state amplitude, as well as the 4th harmonic steady-state amplitude. The corresponding noise floors are indicated by the dotted lines.

Besides a different order of magnitude and noise floor, the peak-magnitude of the even response and the amplitude of the 2nd harmonic show common properties, like a broad best frequency region and a high-frequency fall-off. The cut-off frequencies and frequency limits (vertical dashed lines) are in reasonable agreement. The main difference is at the lowest frequencies, where the 2nd harmonic steady-state values do not show the plateau seen in the peak-magnitudes of the even harmonic response (discussed in next section). At these

frequencies, there is a significant 4th harmonic, which is however not large enough in amplitude to explain the plateau. Note that the absolute value of the peak-magnitude of the even harmonic response and its corresponding baseline are much higher than those of the steady-state amplitude of the 2nd harmonic. The higher peak-magnitude simply reflects the larger amplitude of the neural response at its onset; the higher noise floor reflects the lack of temporal integration in the extraction method, which reduces the noise in the steady-state measurement. That, despite these differences, the upper limits reasonably agree between the two methods suggests that they reflect actual neural limits rather than limitations in measurement.

Short-time Fourier transformation. We already ruled out that the low frequency plateau reflects higher even harmonics (brown trace, Fig. 13). Close scrutiny of responses reveals a CAP-related component which persists despite the Chebyshev FIR filtering (see METHODS), but only for low frequency probe tones. This is illustrated in Fig. 14A for unmasked probe responses for the two polarities (cf. Fig. 2Aa); in Fig. 14B for unfiltered adapted components (cf. Fig. 2Ac); in Fig. 14C,D, and E for cap filtered (cf. Fig. 2Ad), odd harmonic (difference, as in Fig. 2Bd), and even harmonic (sum, as in Fig. 2Cd) adapted components using the response at 300 Hz also shown in Fig. 13A. The time courses of the amplitudes of the signals in C, D, and E (AN-dNP) are represented by Panel F and show a much larger peak-amplitude than visually expected from the alternating components in the signals. The blue circle indicates a remaining CAP component: some of the CAP spectral components leak through the filter causing a residual CAP (Fig. 13C). This leak through, which was expected, starts to occur as probe frequencies decreases below 600 Hz (its 2nd harmonic is equal to 1.2 kHz), consistent with the upper spectral boundary (~1 kHz) of CAPs. Another residue however was more unexpected and is indicated by the red ellipse in Fig. 14 in the odd harmonic response. One of the assumed properties, and advantage, of the subtracted (polarity dependent) response is that polarity alternation followed by subtraction removes the CAP. Clearly this is not completely the case here: the subtraction leaves a spike-shaped residue that is much shorter than the CAP itself and with a higher spectral content than the probe frequency. The residue is due to a small timing difference between the CAP to opposite polarities, here ~0.52 ms and illustrated with the vertical lines and arrows in Fig. 14A.

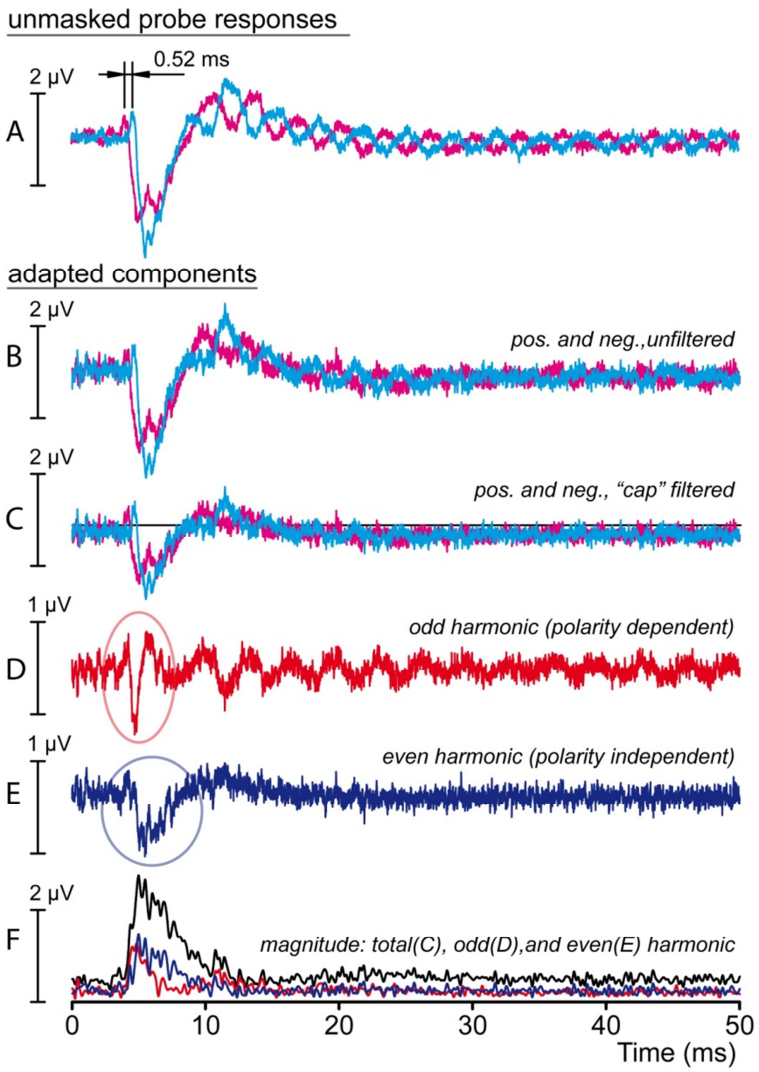


Figure 14. An example of the CAP-related response in the 1st and 2nd harmonic of a low probe frequency.

To cope with the CAP residue and to additionally improve the SNR, we reanalyzed the data using a Gabor transformation, which is a special case of a short-time Fourier transform (STFT) with a Gaussian truncated ($\alpha = 2.5$) window function. The STFT parameters were chosen to have a good SNR performance over the entire frequency range. Fig. 15 shows an example of such STFT for a probe frequency of 1 kHz. Fig. 16 repeats the traces shown in Fig. 13 but omitting the 4th steady-state harmonic. Also shown are the peak magnitudes of the even harmonic response reprocessed with the STFT, shown as the blue solid curves. Comparison with the cyan curves (taken from Fig. 13) shows that the STFT results have a higher SNR and that they parallel the earlier analysis except at frequencies below 600 Hz.

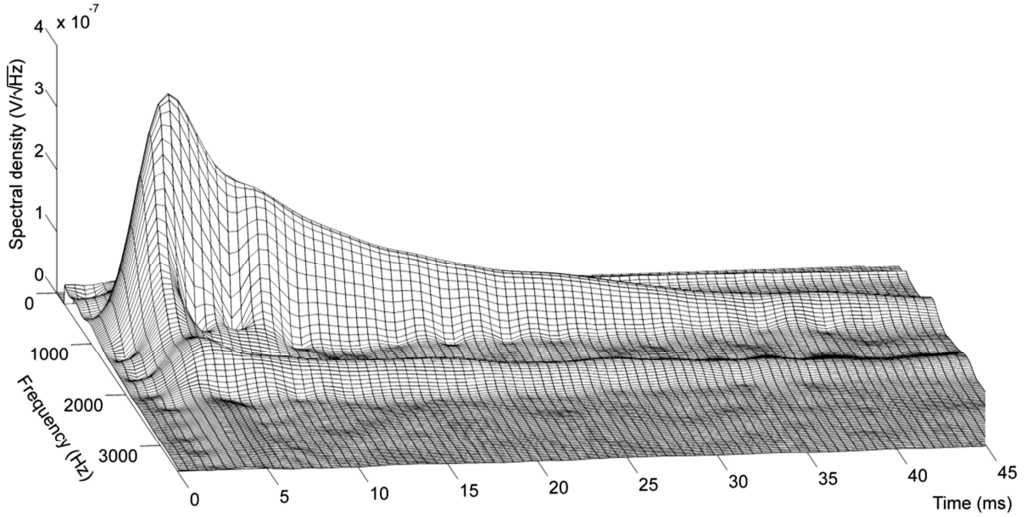


Figure 15. Illustration of a RW-dNP surface spectrogram obtained with a Gabor transformation. The probe frequency was 1 kHz; the signal at 2 kHz is the 2nd harmonic. To mimic the amplitude envelope, the Y-axis is expressed as the square root of the power spectral density (V/√Hz).

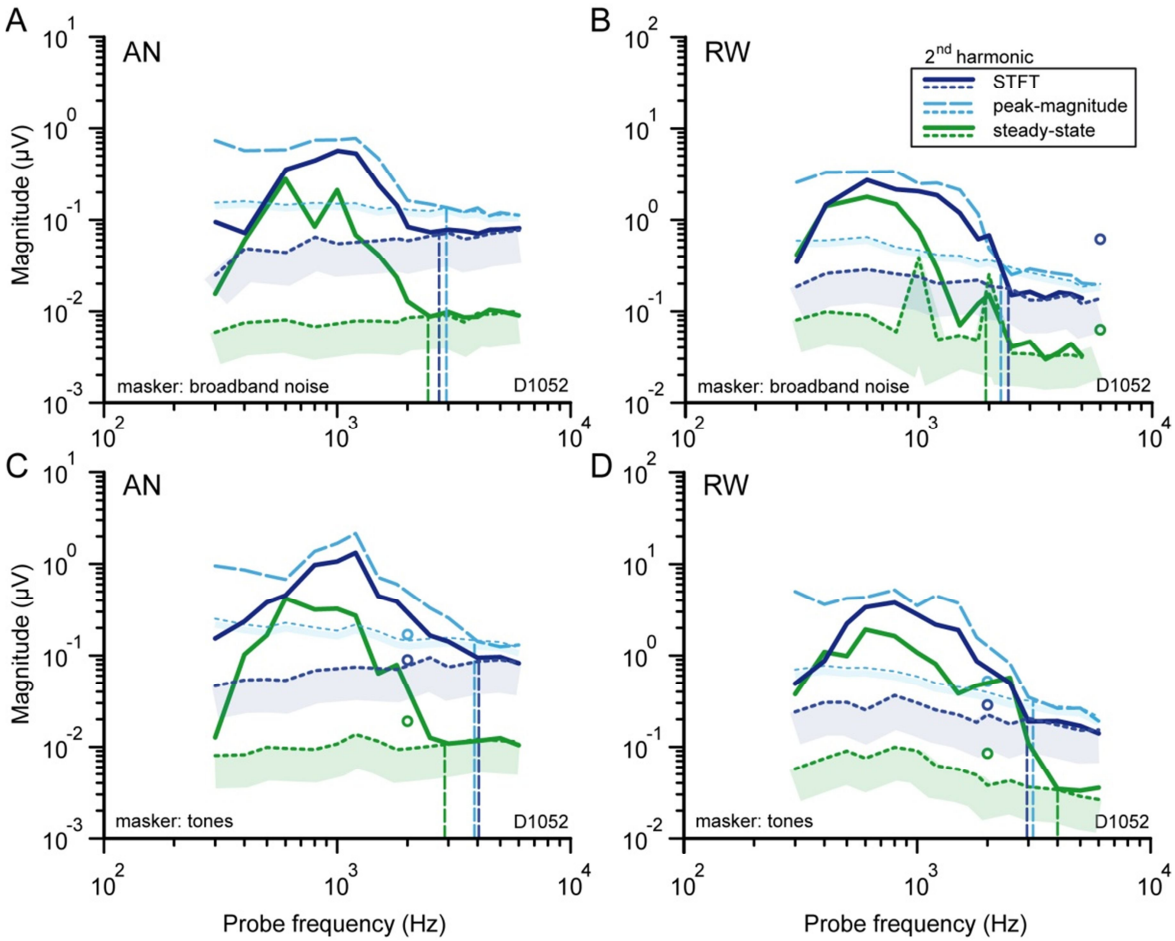


Figure 16. Comparison of the 2nd harmonic reprocessed using the STFT-method (dark blue). Cyan and green: peak-magnitude and steady-state amplitude as calculated in Fig. 13.

Tonal versus noise maskers. In initial measurements we used two types of maskers, a fixed broadband noise masker, to masks neural activity over a broad frequency range, and tonal maskers with the same frequency as the probe. We compared the results for both types of maskers and found no major differences, only some small deviations at the lowest probe frequencies. This is illustrated in Fig. 17A-F for the maximum magnitudes obtained with the STFT method applied to the AN-dNP (A-C) and RW-dNP (D-F).

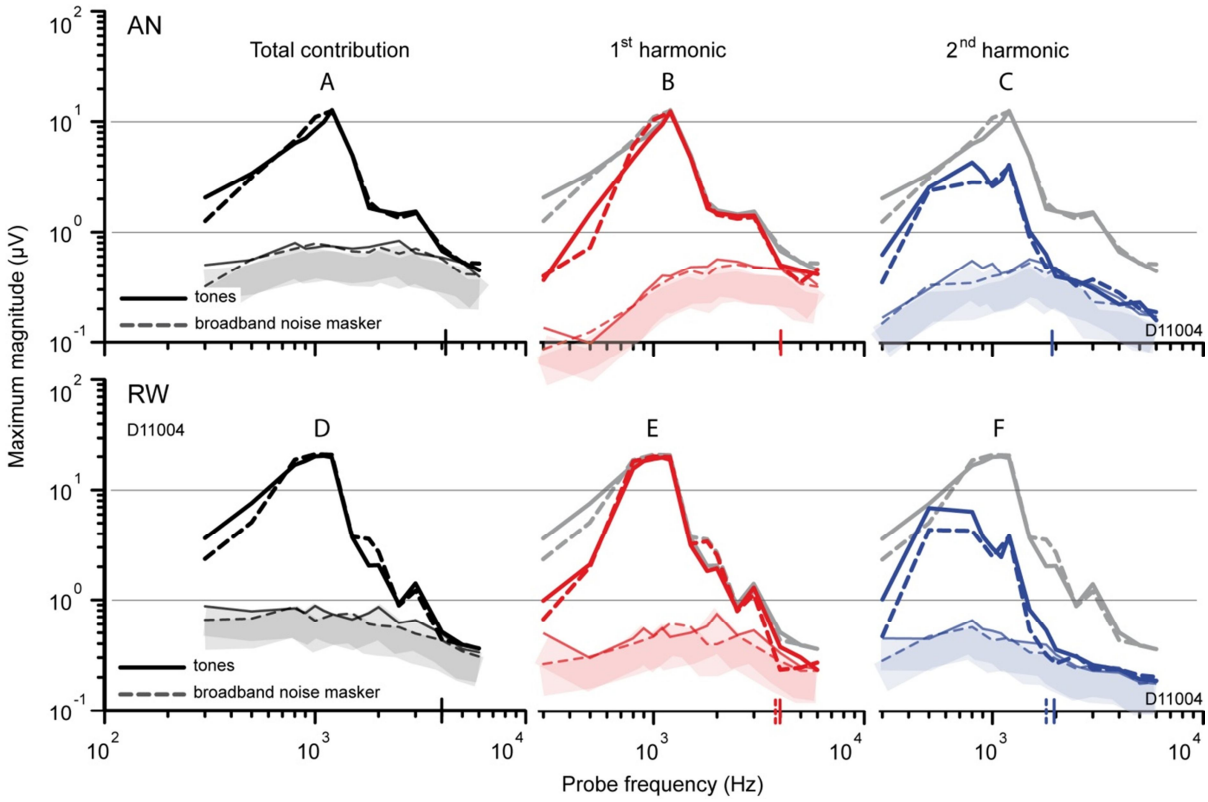


Figure 17. Comparison between the maximum AN-dNP (A-C) and RW-dNP (D-F) amplitudes for tonal and broadband noise maskers using the STFT method. Panels A and D contain maximum amplitudes obtained from all harmonic contributions; and B and E contain the 1st harmonic and the right curves the 2nd harmonic. For visual reference, the curves of A and D are reshown in the background of the others. The masker was a tone with the same frequency as the probe tone. The level for the probes was 50 dB SPL and for the masker 70 dB SPL. The corresponding noise floors are indicated by thin lines and shading. The vertical ticks inside the abscissa indicate the upper limits at which magnitude merge their noise floor.

The analysis is applied to the total harmonic response (A,D); the odd harmonic response (B,E); and the even harmonic response (C-F). Finally, the results for both the noise maskers (dashed lines) and tone maskers (solid lines) are shown. The small differences between tonal and noise masking at the lower frequencies indicate slightly more effective masking by tones; more intense noise maskers would give the same result as tones. Importantly, the upper frequency limit is independent of the masker type - as expected - and was approximately 4 kHz for the total and 1st harmonic contribution, which is somewhat lower than the results from the experiments shown above (Fig. 9-12). The upper limit for the 2nd harmonic is, at ~2 kHz, very consistent with that obtained earlier. The center frequency of the broad peak is also quite consistent for the different curves, and is again ~1.2 kHz. Note that so far we have assumed that the energy of phase-locking was mainly concentrated in the 1st harmonic and to a lesser extent in the 2nd harmonic. This is visually the case for most frequencies as shown by the gray overlays in panels B,C,E, and F of Fig. 17. However, at very low frequencies this is no longer true. For example, at 500 Hz the 2nd harmonic contribution (panel C and F) is larger than the 1st harmonic contribution (panel B and E). At 300 Hz, the total harmonic contribution is clearly higher than this of the sum of the 1st and 2nd harmonics. For this particular frequency we calculated the 3rd and 4th harmonic and found that their contributions were larger than those of the 1st and 2nd harmonics (not shown). Therefore, to obtain accurate amplitudes at very low probe frequencies, higher harmonics, up to the 4th, have to be taken into account. In contrast, to obtain the cut-off frequency or the upper limit of phase-locking the 1st harmonic suffices.

TTX

As a final test of the neural basis of the components that we designate as such, we compared results obtained before and after the application of Tetrodotoxin (TTX) at the RW in one animal. TTX blocks the firing of action potentials in nerve cells, but does not affect the ionic transduction currents through the hair cells. We used the neural blocker to measure the RW-CM in the absence of neural contributions, and to test the effectiveness of our method to cancel the RW-CM.

Effect on CAP. After an initial "normal" measurement session to obtain baseline estimates of the AN-dNP and RW-dNP, we applied 4 µL (10 mM) TTX in the niche of the RW and monitored the CAP responses over a long time. The evolution of the magnitudes of the CAP response is shown in Fig. 18A for 8 probe frequencies ranging from 1 to 10 kHz.

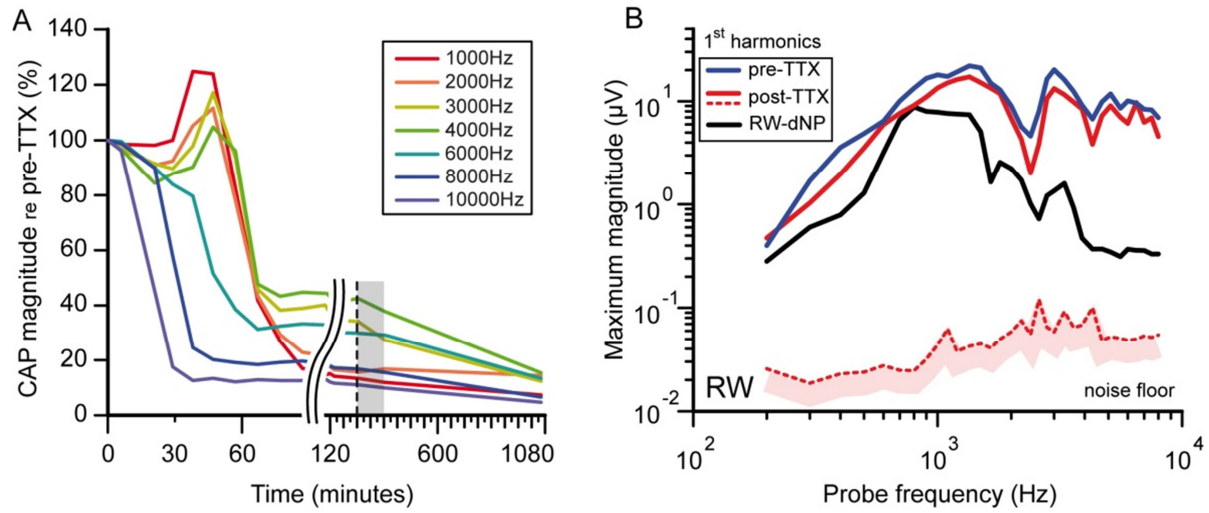


Figure 18. Reduction of CAP but not CM after TTX application. A) CAP reduction as function of time after administration of TTX in the niche of the RW. The parameter is the probe frequency. The dashed line indicates the start of the post-TTX recordings; the gray zone covers the time period of the measurement. B) Comparison between RW-CM magnitudes before and after TTX administration. The pre-TTX curve is obtained at neural onset and this of post-TTX at steady-state using an FFT. For comparison, the corresponding neurophonic of Fig. 19B is included. The corresponding noise floor is indicated by the dotted line.

The CAP at different frequencies is reduced with different onset post-TTX, and these times are consistent with an expanding spatiotemporal diffusion pattern of TTX through the cochlea. The CAP for high frequencies, which excite the most basal receptors, was affected first, followed by the CAP in response to lower frequencies. Remarkably, for frequencies at 3 kHz and below, there was a period of enhancement where the CAPs grew above their initial magnitudes. The CAP then declined simultaneously at these different frequencies. The criterion to start acquisition of a new set of neurophonic measurements was a stable CAP reduction for all measured frequencies. The time of start was 4 hours after the administration of the TTX as indicated by the dashed line. Note that at the time of neurophonic measurement (gray zone in Fig. 18A), there is a remaining residue which is highest for mid-frequencies. This residue is thought to arise from the excitatory post-synaptic potentials of the afferent dendrites under the inner hair cells (Dolan et al., 1989). After an additional time period of 2 hours over which the neurophonic measurements were completed, the CAP was measured again: some small additional magnitude decline was measured at mid-frequencies. A further decrease was noted at a final repeated measurement 18 hours after TTX application.

Effect on CM. The decrease in CAP responses does not simply reflect a loss of cochlear sensitivity. Fig. 18B illustrates the effect of TTX on the RW-CM values. As illustrated earlier (Fig. 6), the probe responses contain a short initial segment over which only

a CM response is present, before the onset of the neural response. Fig. 18B (blue line) shows the RW-CM (1st harmonic) amplitudes measured at neural onset of masked responses, before administration of TTX. At low frequencies the amplitudes show a high-pass behavior with slope of 40 dB/decade. Above 1.2 kHz the spectrum has little overall slope but shows a spectral fine structure with distinct valleys and peaks. The red line shows measurements obtained after administration of TTX, during the time period indicated with shading in Fig. 18A. The noise floor of the RW-CM after TTX is also included. Measurements of 1st harmonic at steady-state and at response onset are nearly identical (only the steady-state value is shown in Fig. 18B). Moreover, these values are very close to RW-CM values measured before application of TTX and decreased with only a few dB. The similarity of these results is consistent with a purely neural effect of TTX, and moreover testifies to the validity of the onset CM as a reliable measure not only of the onset but also of the steady-state hair cell contribution. The small amplitude decrease and the presence of the peaks and valleys in the fine structure are consistent with previous studies (Henry, 1995; He et al., 2012) if it is taken into account that these studies did not attempt to differentiate the neural and non-neural contributions measured at the RW. Taken together, the results indicate that the RW-CM is of the same origin before and after TTX application, and that it contains no or no significant neural activity.

Effect on neurophonic. The magnitudes of the AN-dNP (left column) and RW-dNP (right columns) before (upper row) and after (lower row) TTX administration are shown in Fig. 19. Comparisons between the upper and lower rows shows that TTX has a large suppressive effect on the neurophonic at both locations. The total neurophonic of the AN was reduced by more than 20 dB and became indistinguishable from the background noise. For the RW, the reduction was 26 dB, but some components in the total neurophonic, more specific the 1st harmonic (Fig. 19 , red) remained visible between 0.4 and 2 kHz. Since this frequency range corresponds to the broad peak of the neurophonic (panel B), this residue is almost certainly of neural origin and remains due to an incomplete neural block, as supported by the remaining CAP (Fig. 18). Unfortunately, this could not be verified with the AN because of its relatively higher noise floor.

Remarkably, the noise floor of the RW decreased much more than that of the AN and even reshaped. The decrease is indicated in Fig. 18D by the gray arrows between the total neurophonic before (gray dotted curve) and after (black dotted curve) TTX. The decrease in noise floor for the RW was ~11 dB and for the AN (Fig. 18C) it was only ~1.3 dB. In contrast to the AN, the noise floor of the RW was reshaped from a broad band-pass shape to a straight line. The slope of this straight line is +7 dB/decade, which is close to the slope

(+10 dB/decade) imposed by the frequency proportional filter in our method (verified by simulations). The change in spectrum of the RW noise floor to a flat slope close to 10 dB/decade suggests that the background noise of the cochlea became whiter, which is a good indication that the (spontaneous) neural background activity is quenched (Dolan et al., 1990; McMahon and Patuzzi, 2002; Patuzzi et al., 2004; Searchfield et al., 2004). It is unclear why the noise floor of the AN did not reduce with the same amount as for the RW. One possibility is that the background noise at the AN only partially reflects spontaneous neural AN activity. The measurements in this experiment were recorded with respect to the nape of the neck and not longitudinal over the nerve as in other experiments. Since the AN is not electrically isolated from neighboring structures such as the vestibular nerve, the noise floor is perhaps dominated by generators other than AN fibers.

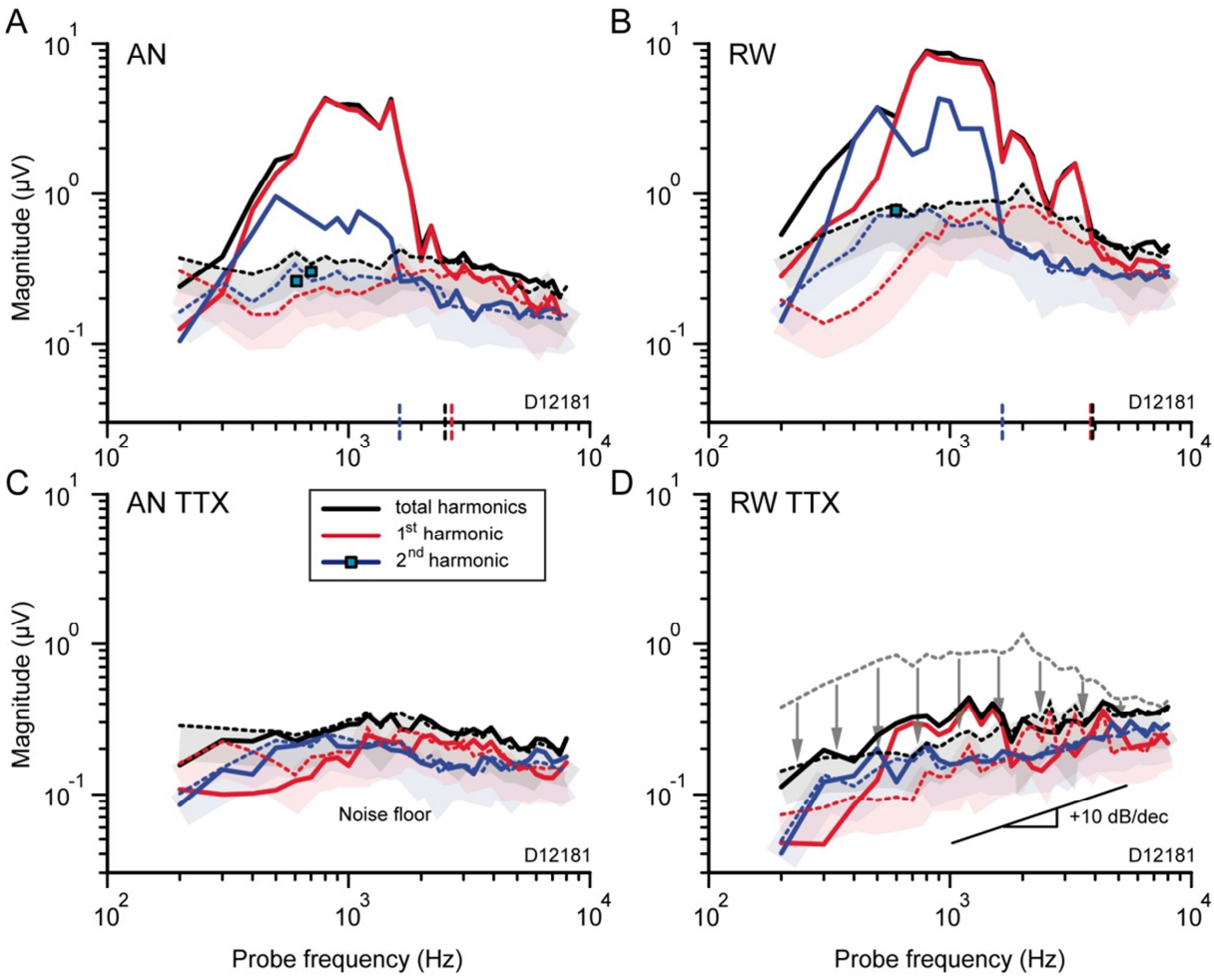


Figure 19. The magnitude of the AN-dNP and RW-dNP before (A,B) and after (C,D) the administration of TTX, for the AN (A,C) and RW (B,D). The corresponding noise floors are indicated by the dotted lines. The slope of +10 dB/decade in D indicates the slope imposed by the frequency proportional filter in the STFT method. The gray arrows in panel D indicate the reduction of the (averaged) background noise at the RW as a result of the administration of TTX.

An interesting feature of the 1st harmonic of the RW-NP is that it shows several similarities to the RW-CM. Fig. 18B (black line) shows this component: like the RW-CM (blue and red curve), it has a low frequency high-pass slope up to ~1 kHz, and a fine structure at higher frequencies that partially resembles that of the CM. We will make use of this relationship in DISCUSSION.

Steady state and 2nd Harmonic. We already mentioned (RESULTS, Quantification of the neurophonic) that the 2nd harmonic measured in the steady state portion of stimulus polarity independent (summed) response (Fig. 5Ca) has been used as a measure of neural phase-locking (Lichtenhan, 2013). Having shown that the RW-CM after TTX contains little neural contribution (Fig. 19), we are in an excellent position to examine whether this steady-state 2nd harmonic can be used to find the frequency limit of phase-locking. The issue is whether that component reflects neural phase-locking or not.

Fig. 20A (blue) shows steady-state 2nd harmonic measured at the RW before TTX application. Note that it is reasonably similar to the 2nd harmonic of the RW-dNP illustrated in Fig. 19B, but that it shows several peaks above 2 kHz which are not present in Fig. 19B. The response and noise floor after TTX application (Fig. 20A, red) are clearly reduced (up to 30 dB) for frequencies < 2 kHz, indicating a strong contribution of neural activity to the steady-state 2nd harmonic at these frequencies. In contrast, the probe responses (with TTX) >2 kHz show only a small reduction in amplitude or they are bounded by the noise floor (dotted line). This, combined with the observation that the peaks > 2 kHz are not seen in the 2nd harmonic of the RW-dNP, strongly suggests that these contributions are not neural. Similar peaks are seen in the 1st harmonic of the RW-CM (Fig. 18B), indicating that the responses > 2 kHz are likely dominated by CM. Small CM reductions after TTX application have been reported before (Henry, 1995; He et al., 2012). In conclusion, the 2nd harmonic steady-state response at high frequencies is not suited to measure neural phase-locking.

TTX did not completely block neural responses (see RESULTS, Fig. 19D), so that it is possible that the remaining steady-state 2nd harmonic (Fig. 20A, redline) is neural. To investigate whether or not responses at frequencies <2 kHz are also dominated by non-neural responses, we compared (Fig. 20B) the amplitudes of the steady state 2nd harmonic after TTX (red trace, same as in Fig. 20A) with the 2nd harmonic from the adapted component (RW-dNP) using the STFT method in this study (green trace, same as blue trace in Fig. 19D).

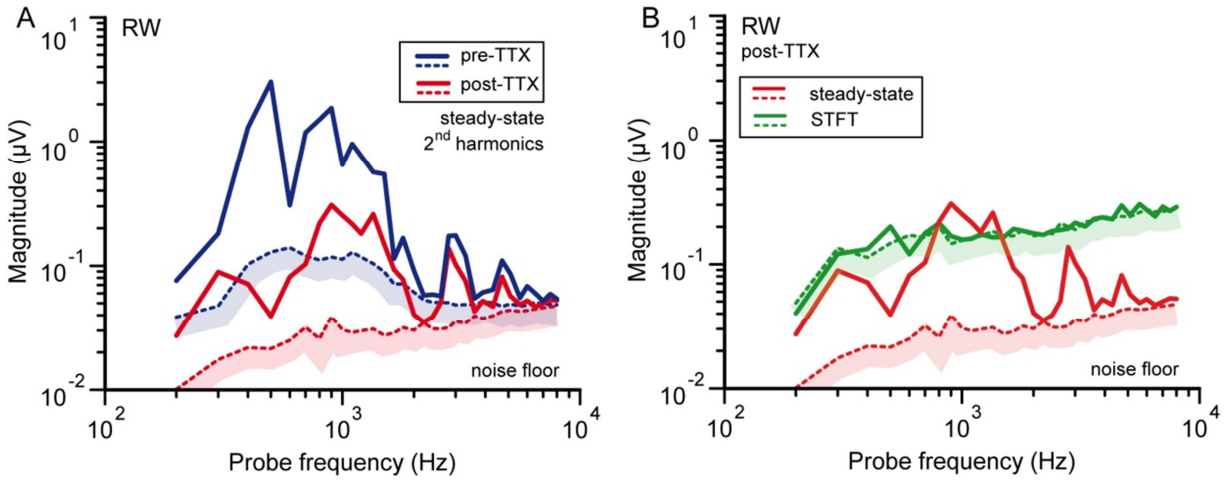


Figure 20. Comparison of 2nd harmonic amplitudes before and after TTX administration at the RW. The amplitudes were obtained from the steady-state part of the probe response using an FFT with long time window.

The noise floors have a general slope of 10 dB/decade and are separated by 18 dB these features are, as expected, from the method of calculation (slope) and the longer responses available for the steady-state measurement (lower noise floor). There is no clear evidence for a response over the complete probe range in the RW-dNP, not even in region between 800 and 1300 Hz where the steady-state amplitude is larger. (Recall that the amplitude of the STFT is considerably larger than that of the steady-state, Fig. 16B and D, so that any response in the STFT would be expected to be much larger than in the steady state). The conclusion is that the steady-state 2nd harmonic is contaminated by CM, while such contamination is absent in the RW-dNP.

To summarize, the steady state 2nd harmonic signal at the RW is not suitable for determination of the upper frequency limit of phase-locking, because of the presence of a non-maskable component which most likely arises from the CM.

Population analysis

In previous sections, we focused primarily on the method and its verification. In this section we combine the entire data set obtained with the STFT-method and determine an overall upper frequency limit of phase-locking. Fig. 21 shows the single data points of 6 sessions for the AN (panel A) and RW (panel B) together with the population trendlines (MATLAB, *RLOESS*, span: 0.35) of the total dNP (black), its 1st (red) and 2nd (blue) harmonic. For both measurement locations, the different trendlines show the typical band-pass

shape with a center frequency around 1 kHz. At this center frequency and above, the 1st harmonic is the dominant component; below ~500 Hz, the 2nd harmonic is clearly present with a contribution comparable to, or even larger, than that of the 1st harmonic. Not shown here, but at the lowest frequencies (200 and 300 Hz) there is a significant contribution of the 3nd and 4th higher harmonics. We remark that responses to probe frequencies above 7 kHz were measured in only one experiment, which biased the trendlines and made the curves only useful up to 6.5 kHz. This is noticeable in Fig. 21 as the slight upturn above 6 kHz.

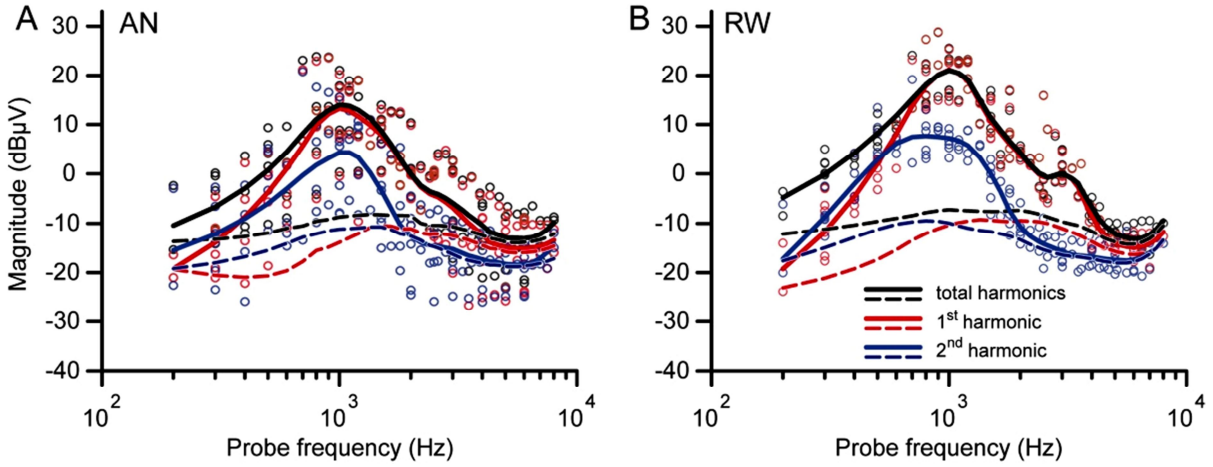


Figure 21. General mean trends for the different harmonics of the neurophonic using the STFT method for the AN and RW. The trends (MATLAB, *RLOESS*, span: 0.35) were obtained from 6 measurement sessions. The corresponding noise floors are indicated by dashed lines. Note that magnitude is in dB re 1 μV_{RMS}.

The determination of an upper frequency limit is only useful with a baseline. Therefore, trendlines of the noise floors were calculated (MATLAB, *RLOESS*, span: 0.35) and included in Fig. 21 (dashed lines; single data points not shown). The trendlines for the 2nd harmonics touch clearly their noise floor, while this is not the case for the 1st and total harmonics. The trendlines of the neurophonic start to bend towards their noise floors but do not touch and remain ~1 dB above them. This phenomenon can probably be assigned to a bias of a few dNP-curves that have not been settled yet at the highest frequencies, i.e. Fig. 11B and Fig. 12B. Whatever its cause, the fact that the signal remains at a constant level of ~1 dB above the noise floor indicates that at that point the noise floor would have overtaken the response. To determine an upper cutoff frequency we first raised the noise floors (by 0.2-1.5 dB) till they merged with the high-frequency part of their neural trendlines, where both curves run parallel. Next, we subtracted these noise floors from their trendlines to remove their influence; the assumption here is that the background noise and phase-locked

signals are not correlated. The result of this subtraction is depicted in Fig. 22 by the solid lines which have a decreasing slope extending up to -infinit (= intersection between the noise floor and the original trend); the curves were truncated at -30 dB and are called the compensated trendlines. We then defined the upper frequency limit at 10 dB below the intersection of raised noise floor and compensated trendlines; this 10 dB point is indicated by the small horizontal lines. At the frequency at which the noise floor and compensated trendline intersect, illustrated for one case with a white dot (Fig. 22B), the SNR for the original trendline is theoretically 3 dB. The choice of 10 dB corresponds to an SNR of ~1.7 dB. Using our new definition we obtained a mean detectable upper frequency limit of phase-locking for the RW at 5.4 ±0.6 (SEM) kHz and 4.2 ±0.6 kHz for the AN. The limit for the 1st harmonic for the RW was at 4.7 ±0.5 kHz and at 4.2 ±0.4 kHz for the AN. The 2nd harmonic had in general a much lower limit, 2.2 ±0.2 kHz for the RW and 1.8 ±0.1 kHz for the AN. Note that the SEM is calculated as the product of the estimated upper frequency limit of the population with the relative SEM from the upper frequency limits of the original neurophonic curves.

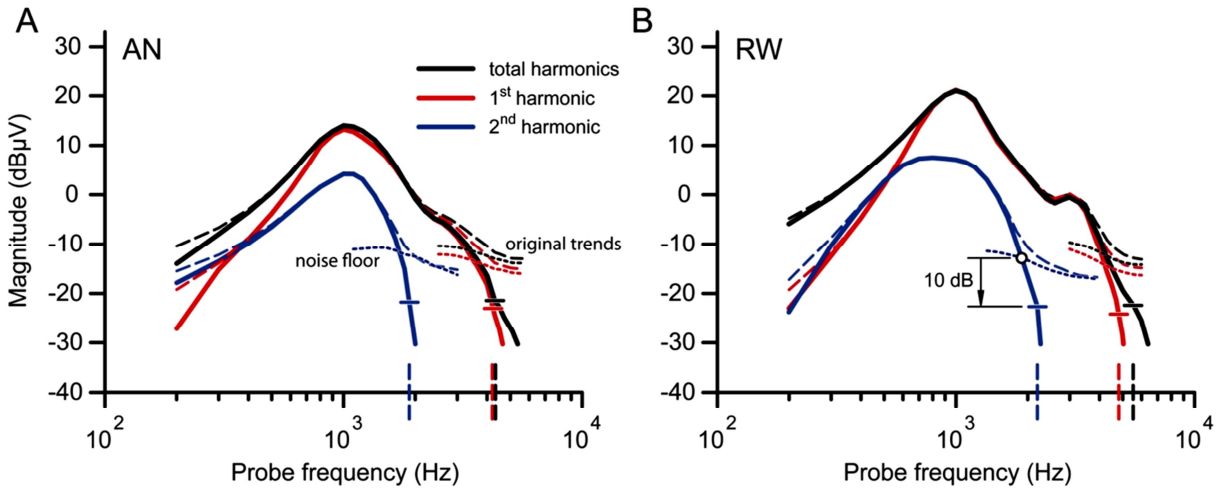


Figure 22. Similar to Fig. 21 (dashed lines), but compensated for the noise floor (solid lines). The upper frequency limits are obtained at the intersection of the compensated trendlines with the horizontal small lines, which are the magnitude obtained at 10 dB below the intersection of the noise floor with the compensated curve. The corresponding noise floors are indicated by dotted lines. The upper frequency limit is indicated with vertical dashed lines on the abscissa.

Probe Level dependence. In this study we did not systematically vary probe level. Nevertheless we have some data regarding its effect on the neurophonic. It is known that the magnitude of the neurophonic tends to increase with probe level (Snyder and Schreiner, 1984) and which can enhance the SNR and perhaps improve determination of the upper frequency limit. Fig. 23 shows the influence of increasing probe level (30, 35… 65 dB SPL; from red to blue color) on the time-course of the RW-dNP (panel A, cf. Fig. 5Ad) and the masked response (panel B, cf. Fig. 5Ab) at 800 Hz. Both types of signals show increasing amplitude with probe level, but with distinctly different behavior and timing. This is illustrated in Fig. 24, where the amplitudes (Fig. 24A) and the delays of the fine structure (Fig. 24B) of two frequencies (0.8 and 1.2 kHz) are plotted as a function of probe level. The amplitude of the estimated RW-CM (green dashed lines, Fig. 24A) shows a monotonically increasing dependence on probe level; moreover, the amplitude has a linear (1 dB/dB SPL) relationship with intensity and is consistent with that reported in a previous study (figure 16, post-TTX, Snyder and Schreiner, 1984). In the contrary, the peak-amplitude of the RW-dNP (red lines) has no linear relationship. For small levels, the slope is much steeper (800 Hz: 2 dB/dB SPL) than 1 dB/dB (linearity) and becomes shallower (<1 dB/dB) with increasing level. The phase of the fine structure at the onset of the masked response (Fig. 23B) is invariant with level. In contrast, the phase of the fine structure of the RW-dNP (Fig. 23A) advances with level, similar to well-known effect on the CAP (Peake and Kiang, 1962; Eggermont, 1976b; Salvi et al., 1979; Prijs, 1980; Moller, 1986; Versnel et al., 1992a). This phase delay is shown in Fig. 24B as a function of probe level, relative to the phase of the RW at 65 dB. At 800 Hz (solid blue), the phase shift has a monotonous course with a slope of -330 μs/dB SPL at 800 Hz and -195 μs/dB SPL at 1200 Hz. The AN-dNP had a similar dependency on probe level for amplitude and phase as the RW-dNP, but with a slopes (Fig. 24, red) which were, remarkably, less steep (-200 μs/dB SPL at 800 Hz, -135 μs/dB SPL at 1200 Hz,) than that of the RW-dNP. These observations suggest that effects of probe level may be interesting to further tease out neural contributions from the mass-potentials recorded at RW and AN and to gain a better understanding of the generators involved.

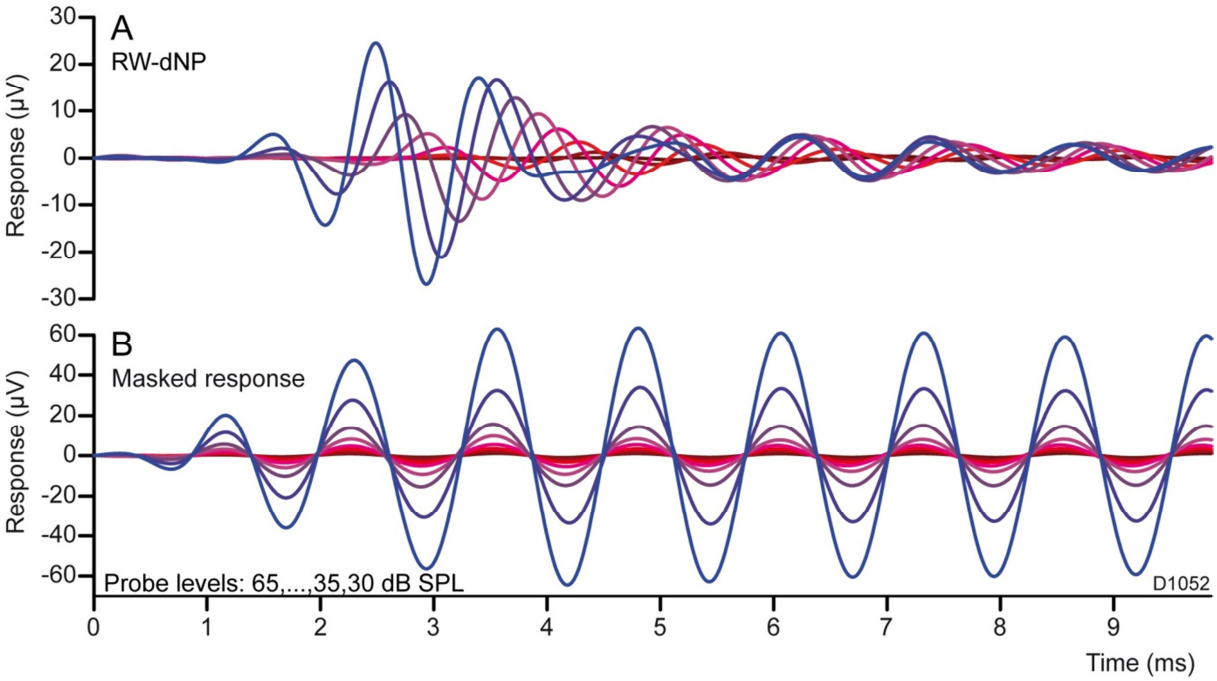


Figure 23. Influence of probe level on the time-course of the RW-dNP (A) and masked response (B) for a probe frequency of 800 Hz. The frequency of the tonal masker was the same as this for the probe; the masker level was 65 dB SPL. The hue indicates the probe level, dark red is lowest level, blue is highest.

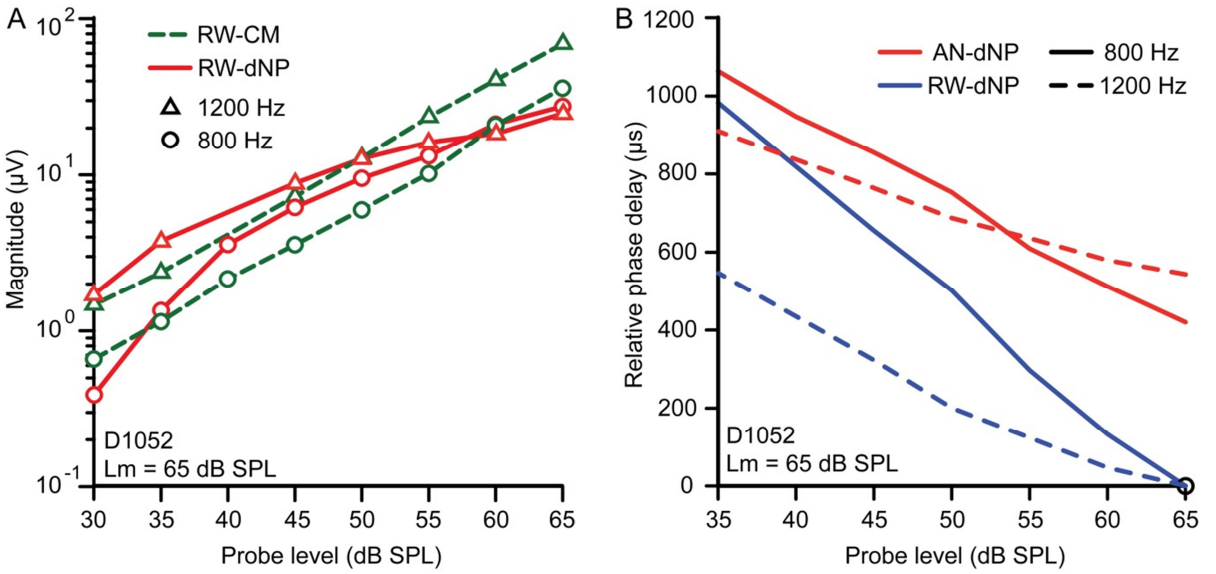


Figure 24. Peak-magnitudes (A) and relative delays of the fine structure (B) as a function of probe level for two frequencies (0.8 and 1.2 kHz). (A) Trendlines of the peak-magnitudes of the RW-dNP are shown in red; trendlines of the magnitudes of the RW-CM are shown in green. (B) Trendlines for the RW-dNP (blue) and AN-dNP (red) delays relative to the phase of the RW at 65 dB. The delay is expressed in µs.

DISCUSSION

Human behavioral results have been interpreted as indicating a higher limit of neural phase-locking than the limits known in mammal species commonly studied in the laboratory (e.g., Heinz et al., 2001; Moore, 2008). The true extent of neural phase-locking in the AN in human is unknown. We have recently provided arguments against a higher limit in humans (Joris and Verschooten, 2013), but physiological measurements are needed to obtain direct evidence. The aim of the present study was to find a method that can extract a measure for the neurophonic from the mass potentials at the cochlear RW, since such a method might be applicable to human volunteers with normal hearing. To achieve this goal, different techniques were combined. The main challenge is to separate a phase-locked signal generated by the hair cells, usually referred to as the CM, from the neural phase-locked signal, particularly because the CM extends to much higher frequencies than neural phase-locking. To achieve this separation, we used signal averaging, forward masking and filtering. The filtering included non-causal linear FIR filtering, non-linear de-noising (wavelets), Hilbert transformation, spectral analysis (FFT, STFT), and odd/even harmonic separation using stimulus polarity reversals. The technique was first assessed at the AN, then applied at the RW, and was validated using TTX. Previous authors have examined whether limits of neural phase-locking may be gleaned from mass-potential recordings. Snyder and Schreiner (1984) reported neurophonic recordings in cat that suggested neural phase-locking exceeding the generally quoted limit of ~5 kHz; on the other hand Lichtenhan et al. (2013) expressed skepticism in the ability of such recordings to be a valuable tool at high frequencies.

Responses recorded on the AN.

A logical place to first look for neural phase-locked mass potentials is the AN. This site is also potentially accessible in human subjects undergoing neurosurgery (Moller and Jho, 1989). The averaged probe and masked probe responses recorded at the AN (e.g., Fig. 2Aa, Ab) showed similarities with those of previous studies (Weinberger et al., 1970; Snyder and Schreiner, 1985; Chimento and Schreiner, 1990). The probe responses showed a polarity independent CAP and an alternating component that was phase-locked to the stimulus; both components were strongly reduced by a forward tonal or noise masker. We made significant efforts to isolate a signal and metric that capture a purely neural contribution to the alternating potentials measured - the dNP amplitude - and validate the metric with several observations throughout the paper. We showed that the phase-locked part of the adapted component (AN-dNP) consists of neural contributions without apparent contamination by CM. The strongest

arguments are 1) that this component completely decays due to the recovery of masking (Fig. 2Ad) with a time constant consistent to that of Chimento and Schreiner (1990); 2) the clear onset time lag re. CM (Fig. 6), associated with a delay from the afferent synapses and some additional axonal propagation (Palmer and Russell, 1986; Ruggero and Rich, 1987) the nearly complete disappearance of the AN-dNP after TTX application (Fig. 19C). In combination with several internal consistency checks (Figs. 3,6,7,8,18), these arguments provide a compelling case that the AN-dNP captures neural phase-locking. Our recordings also clearly indicate that the CM can make a sizable contribution at the AN (the AN-CM), which grows (relative to the neural contribution) with frequency. We discuss the apparent conflict between our finding of an upper frequency (4.7 kHz at RW) which is rather close to that reported for single-unit AN fibers of cat (~5 kHz), and the results of Snyder and Schreiner (1984) of a neural component at much higher frequencies.

As is clear from many figures reported here, the AN-NP is large near 1 kHz, which is the frequency region where most studies (including ours) parametrically assess the recording method. It is important, however, to keep in mind that the neurophonic at higher frequencies is strongly attenuated, due to the steep low-pass characteristic of single-fiber neural phase-locking (Weiss and Rose, 1988a, b; Kidd and Weiss, 1990), with a slope of at least -40 dB/decade (see below, Fig. 26), while this is not the case for the CM (Henry, 1995; He et al., 2012). It follows that the CM will inevitably dominate above a certain frequency. Fig. 25 shows the trend of the magnitudes of the AN-CM and AN-dNP obtained in this study. The measurements were separated in two groups according to recording method; recordings measured differentially along the nerve (solid lines) and those measured 'monopolar' with respect to the nape of the neck (dashed lines). The AN-CM trends (green curves) were similar for the two methods: they increases up to 0.8-1 kHz, thereafter they decline with a roll-off of -20 dB/decade and finally they intersect the neurophonic at a frequency of 3 kHz for the 'monopolar' recordings and at 4 kHz for the differential recordings. Beyond these frequencies, the AN-CM is larger than the AN-dNP. The AN-dNP trends (red lines) also show a band-pass trend centered near 1 kHz, but note that their low-pass slope has a distinctly steeper roll-off; it is roughly -50 dB/decade.

This contrasts with the roll-offs of transfer functions in the study of Snyder and Schreiner (1984; their Figure 12), obtained with differential recordings using electrodes placed transverse across the AN. The higher probe levels of 70 and 80 dB SPL in that study produced a much less steep roll-off of -20 dB/decade compared to that of a probe level of 60 dB SPL (-40 dB/decade) and to that obtained in this study (-50 dB/decade); but the -20 dB/decade roll-off is similar to that for the AN-CM in our study. Moreover, the upper

frequency limits of the transfer functions of the highest levels were not determinable and were much higher than 10 kHz; in contrast, the lowest level had a limit between 3 and 4 kHz and is comparable to the limit obtained in this study. We surmise that the transfer functions in Snyder and Schreiner (1984) are contaminated by stray-CM, and that the degree of contamination depends on the recording method, the probe frequency and the level. The latter is also supported by the fact that the contribution of the CM becomes relatively more important for higher probe levels (Snyder and Schreiner, 1984), as also shown in Fig. 24B.

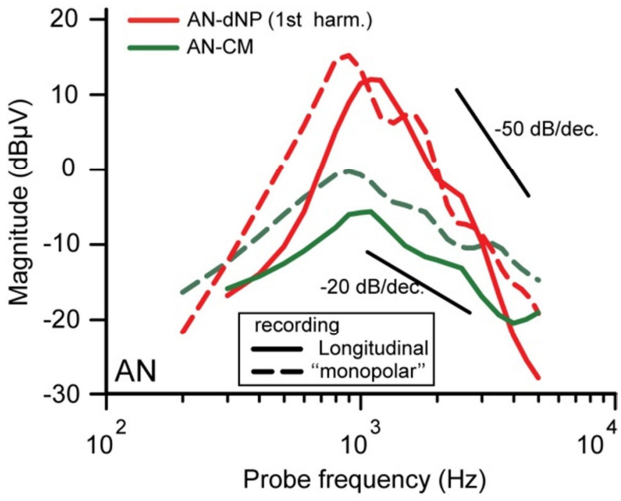


Figure 25. Mean trends of the AN-CM and AN-dNP for two different recording configurations. In the longitudinal configuration the electrodes are placed along the AN; the positive electrode was positioned near the internal auditory meatus and the reference electrode at the merging of the AN with the CN. These results are from 2 measurement sessions in 2 different cats. In the "monopolar" method the positive electrode was positioned on the nerve and the reference electrode in the nape of the neck. The results are also from 2 measurement sessions in 2 different cats. The AN-dNP data were obtained with the STFT method; the AN-CM data were obtained from the masked signal at neural onset. The black lines indicate roll-offs for reference.

Responses recorded on the RW.

The most important conclusion from the present study is that there is a sizable neural phase-locked contribution that can be measured at the RW: the RW-dNP. The presence of such a contribution had been reported earlier (Henry, 1995; Lichtenhan et al., 2013). We show here that it is strikingly similar in many, but not all, properties to the AN-dNP. The main differences were a typically *larger* magnitude at the RW (compare Fig. 7B and D; Fig. 8) and a small time lead at the RW (Fig. 7B and D; Fig. 8). We will return to the time differences in the next section (Fig. 24) but a full treatment is beyond the scope of the present study.

The RW-dNP and AN-dNP share properties with the CAP and with single AN fibers. First, the magnitude of the dNP has a linear relationship with masker level (in dB SPL) over a dynamic range of 20 a 30 dB and this up to about 5 a 10 dB above the probe level (Fig. 7), similar to CAP masking intensity curves (Figure 5, Verschooten et al., 2012) and single AN

fibers (Figure 9 and 12, Harris and Dallos, 1979). Second, the time of onset and the phase of the neurophonic decreased with increasing probe level (Snyder and Schreiner, 1984) (see also Fig. 23 and Fig. 24), as is also the case for the CAP (Verschooten et al., 2012) and, with some restrictions, for single AN fibers (Allen, 1983; Carney and Yin, 1988; Palmer and Shackleton, 2009). On the other hand, the masker level (Fig. 7) had no influence on the time-related characteristics. Third, the onset-time of the neurophonic showed a decreasing relationship with frequency (Fig. 8), as is also the case for the CAP and for single AN fibers for tones at CF (Antoli-Candela and Kiang, 1978).

Use of the 2nd harmonic to measure neural phase-locking.

As mentioned in the INTRODUCTION, neural rectification introduces higher harmonics, which can be used to disambiguate CM and NP. More specifically, a second harmonic component is present in the summed responses to opposite polarities. Is this polarity independent response a suitable signal for the determination of the frequency limit of phase-locking? It was recently studied by Lichtenhan et al. (2013), who refer to this response as the ANOW (auditory nerve overlapped waveform) and who stated it cannot be used to find the upper phase locking limit of the (human) auditory system. Their reasoning is that, even though neural responses may be synchronized at e.g. 2 kHz, the effective so-called unit response of each phase-locked action potential has a certain duration (~0.5 ms: Kiang et al., 1976; Prijs, 1986) which smoothes out the phase-locked mass potential. Note, however, that this factor does not imply a hard limit to detect phase-locking at high frequencies, but rather that it imposes a practical limitation. The low-pass filtering due to convolving effect of the unit response strongly reduces response magnitude and thus SNR, but it can in theory be overcome by increasing the measurement time. Still, we find (Fig. 20) that the steady-state 2nd harmonic is not suitable to determine the upper limit of phase-locking, because CM contamination effectively limits detection of very small neural signals at the upper edge of phase-locking.

Relationship to phase-locking in single AN fibers

To finish, we examine how our results relate to the synchronization of single AN fibers. As mentioned before in the RESULTS, a direct comparison between the vector strength (VS, Goldberg and Brown, 1969) of single nerve-fibers in cat (Johnson, 1980; Joris et al., 1994b) and the amplitudes of the neurophonic is hampered by the nature of the neurophonic. First, the neurophonic is an ensemble response from many AN fibers, with their own timing, rate and vector strengths depending on frequency, CF, and stimulus level (Kim

and Molnar, 1979; Palmer and Russell, 1986; Palmer and Shackleton, 2009; Temchin and Ruggero, 2010). Spatiotemporal summation affects the amplitude and phase of the recorded response in ways that are ill-understood due to the complex nature of cochlear mechanics. Second, whereas the VS of AN fibers is based on the *timing* of the action potentials, the neurophonic is an analog waveform composed of summed single unit responses (Kiang et al., 1976; Prijs, 1986; Versnel et al., 1992b) whose waveform has a low-pass filtering effect. Third, the information of the mean activity of the AN fibers, which in VS analysis is used for normalization, is not available in the neurophonic.

In an attempt to compare the neurophonic with single AN fibers we derived a similar kind of synchronization index for the neurophonic based on the RW-CM, which we refer to as neurophonic synchronization. In order to obtain this result, first, the neurophonic had to be compensated for the convolving effect of the unit response waveforms. For this we extracted an average normalized spectrum from 4 unit responses in cat (unit 9, 24, 45 and 43 from Figure 6 in Kiang et al., 1976). Fig. 26A shows a number of frequency dependent curves; the mean unit response (blue); the mean trends of the RW-CM (green); and the mean trends of the 1st harmonic component of the neurophonic (Fig. 22B, red curve). The red dashed line is the 1st harmonic component of the neurophonic, but now compensated for the filtering effect of the waveform of the unit response. The brown line is the neurophonic synchronization for the neurophonic and is calculated as the ratio between the compensated neurophonic (red dashed curve) and that of the RW-CM (green curve). In this procedure, we use the RW-CM as a reference signal for the neural contribution. Only the spectral shape of the resulting curve is of interest: its maximum is put at 0 dB. Finally, the VS (here expressed in dB) for single AN fibers in cat (circles) is shown in black (extracted from Figure 5 of Johnson, 1980).

The neurophonic synchronization (brown) does not show the same smooth course as the VS (black line). The neurophonic synchronization fluctuates between 0 and -10 dB and declines at a higher frequency (3.3 kHz) than the VS (~2.5 kHz). The trends of the neurophonic and RW-CM exhibit a peculiar and similar oscillating feature between 1 and 4 kHz, which is also to a lesser extend present in the neurophonic of the AN (Fig. 25). Our main interest, however, is in the overall slopes of the trends. Idealized curves are shown in Fig. 26B with slope values in dB/decade. The neurophonic synchronization shows a roll-off around 3 kHz with a slope of -60 dB/decade, reasonably consistent with trendline for VS in single fibers. Of course, a full account of the relationship between single-fiber responses and the neurophonic would require more detailed knowledge at many levels.

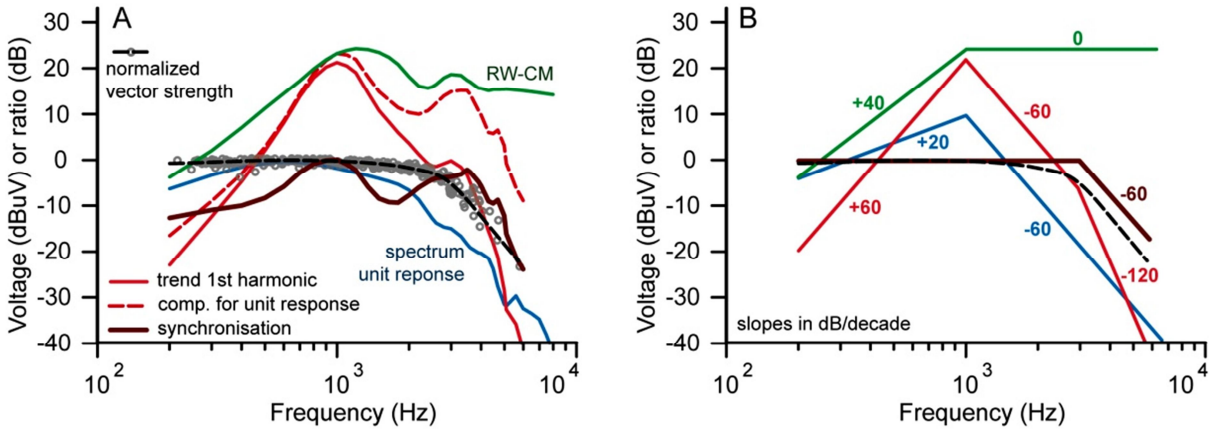


Figure 26. Neurophonic synchronization. (A) Different frequency dependent curves are shown: from a mean unit response (blue) (from 4 unit responses of Kiang et al., 1976), the mean trend of the RW-CM (green), the mean trends of the 1st harmonic component of the neurophonic (Fig. 22B, red curve) and its compensated for the filtering effect of the waveform of the unit response (red dashed curve), the neurophonic synchronization (brown line; calculated as the ratio between the red dashed curve and green curve) and the VS (expressed in dB) for single AN fibers in cat (black dashed line; the single data points indicates as gray circles are extracted from Figure 5 of Johnson, 1980). (B) Same as in A but for idealized curves. The slopes are indicated next to the curves.

CONCLUSIONS

In this study we showed that the potential near the round window has a considerable neural phase-locked contribution that can be isolated and quantified. We found that this component, the RW-neurophonic, is very similar to that measured on the auditory nerve, but with a typical lager amplitude and smaller time lag. The frequency limit of neural phase-locking obtained in cat was 4.7 kHz and very similar to that reported for individual auditory nerve fibers (~ 5 kHz).The method developed in this study has the potential to examine neural phase-locking and the upper frequency limit in normal hearing humans.

REFERENCES

Allen JB (1983) Magnitude and phase-frequency response to single tones in the auditory nerve. The Journal of the Acoustical Society of America 73:2071-2092.

Antoli-Candela EJ, Kiang NYS (1978) Unit activity underlying the N1 potential. In: Evoked Electrical Activity in the Auditory Nervous System (Naunton R, Fernandez C, eds), pp 165-189. New York: Academic Press.

Blackburn CC, Sachs MB (1992) Effects of off-BF tones on responses of chopper units in ventral cochlear nucleus. I. Regularity and temporal adaptation patterns. J Neurophysiol 68:124-143.

Carney LHC, Yin TCT (1988) Temporal Coding of Resonances By Low-Frequency Auditory Nerve Fibers: Single-Fiber Responses and a Population Model. J Neurophysiol 60:1653-1677.

Chimento TC, Schreiner CE (1990) Time course of adaptation and recovery from adaptation in the cat auditory-nerve neurophonic. J Acoust Soc Am 88:857-864.

Dallos P (1973) The auditory periphery. New York: Academic Press.

Dolan DF, Xi L, Nuttall AL (1989) Characterization of an EPSP-like potential recorded remotely from the round window. J Acoust Soc Am 86:2167-2171.

Dolan DF, Nuttall AL, Avinash G (1990) Asynchronous neural activity recorded from the round window. The Journal of the Acoustical Society of America 87:2621-2627.

Dunai L, Hartmann WM (2011) Frequency dependence of the interaural time difference thresholds in human listeners. J Acoust Soc Am 129:2485–2485.

Eggermont JJ (1976a) Electrophysiological study of the normal and pathological human cochlea. II. Neural responses. Rev Laryngol Otol Rhinol (Bord) 97 Suppl:497-506.

Eggermont JJ (1976b) Electrocochleography. In: Handbook of sensory physiology (Keidel WD, Neff WD, eds), pp 625-705. Berlin: Springer.

Goldberg JM, Brown PB (1969) Response of Binaural Neurons of Dog Superior Olivary Complex to Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. J Neurophysiol 22:613-636.

Goldstein MH, Kiang NYS (1958) Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. The Journal of the Acoustical Society of America 30:107-114.

Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. Journal of neurophysiology 42:1083-1107.

He W, Porsov E, Kemp D, Nuttall AL, Ren T (2012) The group delay and suppression pattern of the cochlear microphonic potential recorded at the round window. PLoS One 7:e34356.

Heinz MG, Colburn HS, Carney LHC (2001) Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. Neural Comput 13:2273-2316.

Henry KR (1995) Auditory nerve neurophonic recorded from the round window of the Mongolian gerbil. Hear Res 90:176-184.

Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J Acoust Soc Am 68:1115-1122.

Joris PX, Verschooten E (2013) On the limit of neural phase-locking to fine-structure in humans. Basic Aspects of Hearing: Physiology and Perception Editors: BCJ Moore, RD Patterson, I Winter, RP Carlyon, and HE Gockel (Springer, NY), in press.

Joris PX, Smith PH, Yin TC (1994a) Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. Journal of neurophysiology 71:1037-1051.

Joris PX, Carney LH, Smith PH, Yin TC (1994b) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. Journal of neurophysiology 71:1022-1036.

Kiang NYS, Moxon EC, Kahn AR (1976) The Relationship of Gross Potentials Recorded fron the Cochlea to Single Unit Activity in the Auditory Nerve. In: Electrocochleography (Ruben RJ, Elberling C, Salomon G, eds), pp 95-115. Baltimore: University Park Press.

Kiang NYS, Watanabe T, Thomas EC, Clark LF (1965) Discharge patterns of single fibers in the cat's auditory nerve, 35 Edition. MIT Press, Cambridge: Research Monograph No 35.

Kidd RC, Weiss TF (1990) Mechanisms that degrade timing information in the cochlea. Hear Res 49:181-208.

Kim DO, Molnar CE (1979) A Population Study of Cochlear Nerve Fibers: Comparison of Spatial Distributions of Average-Rate and Phase-Locking Measures of Responses to Single Tones. J Neurophysiol 42:16-30.

Klumpp R, Eady H (1956) Some measurements of interaural time differences thresholds. J Acoust Soc Am 28:859-864.

Kraus N, Bradlow AR, Cheatham MA, Cunningham J, King CD, Koch DB, Nicol TG, McGee TJ, Stein LK, Wright BA (2000) Consequences of neural asynchrony: a case of auditory neuropathy. Journal of the Association for Research in Otolaryngology : JARO 1:33-45.

Lichtenhan JT, Cooper NP, Guinan JJ, Jr. (2013) A new auditory threshold estimation technique for low frequencies: proof of concept. Ear Hear 34:42-51.

McMahon CM, Patuzzi RB (2002) The origin of the 900 Hz spectral peak in spontaneous and sound-evoked round-window electrical activity. Hearing research 173:134-152.

Moller AR (1986) Effect of click spectrum and polarity on round window N1N2 response in the rat. Audiology 25:29-43.

Moller AR, Jho HD (1989) Response from the exposed intracranial human auditory nerve to low-frequency tones: basic characteristics. Hear Res 38:163-175.

Moore BC (2008) The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. J Assoc Res Otolaryngol 9:399-406.

Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair cells. Hear Res 24:1-15.

Palmer AR, Shackleton TM (2009) Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency. Journal of the Association for Research in Otolaryngology : JARO 10:233-250.

Patuzzi RB, Brown DJ, McMahon CM, Halliday AF (2004) Determinants of the spectrum of the neural electrical activity at the round window: transmitter release and neural depolarisation. Hear Res 190:87-108.

Peake WT, Kiang NY (1962) Cochlear responses to condensation and rarefaction clicks. Biophysical journal 2:23-34.

Prijs VF (1980) In: The single-fibre representation in the compound action potential of the auditory nerve, p 91: PhD thesis.

Prijs VF (1986) Single-unit response at the round window of the guinea pig. Hearing research 21:127-133.

Rhode WS, Smith PH (1985) Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers. Hear Res 18:159-168.

Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. Physiological reviews 81:1305-1352.

Rose JE, Brugge JF, Anderson DJ, Hind JE (1967) Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J Neurophysiol 30:769-793.

Rose JE, Brugge JF, Anderson DJ, Hind JE (1968) Patterns of activity in single auditory nerve fibres of the squirrel monkey. In: de Reuck, AVS, Knight, J (Eds), Ciba Foundation Symposium on Hearing Mechanisms in Vertrebrates J&A Churchill, London, pp144–157.

Ruggero MA, Rich NC (1987) Timing of Spikes initiation in Cochlear Afferents: Dependence on Site of Innervation. J Neurophysiol 58:379-403.

Salvi RJ, Henderson D, Hamernik RP (1979) Single auditory nerve fiber and action potential latencies in normal and noise treated chinchillas. Hear Res 1:237-251.

Schiano JL, Trahiotis C, Bernstein LR (1986) Lateralization of low-frequency tones and narrow bands of noise. J Acoust Soc Am 79:1563-1570.

Searchfield GD, Munoz DJ, Thorne PR (2004) Ensemble spontaneous activity in the guinea-pig cochlear nerve. Hear Res 192:23-35.

Snyder RL, Schreiner CE (1984) The auditory neurophonic: basic properties. Hearing research 15:261-280.

Snyder RL, Schreiner CE (1985) Forward masking of the auditory nerve neurophonic (ANN) and the frequency following response (FFR). Hear Res 20:45-62.

Starr A, Sininger Y, Nguyen T, Michalewski HJ, Oba S, Abdala C (2001) Cochlear receptor (microphonic and summating potentials, otoacoustic emissions) and auditory pathway (auditory brain stem potentials) activity in auditory neuropathy. Ear Hear 22:91-99.

Temchin AN, Ruggero MA (2010) Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics. J Assoc Res Otolaryngol 11:297-318.

Verschooten E, Robles L, Kovacic D, Joris PX (2012) Auditory nerve frequency tuning measured with forward-masked compound action potentials. Journal of the Association for Research in Otolaryngology : JARO 13:799-817.

Versnel H, Schoonhoven R, Prijs VF (1992a) Single-fibre and whole-nerve responses to clicks as a function of sound intensity in the guinea pig. Hearing research 59:138-156.

Versnel H, Prijs VF, Schoonhoven R (1992b) Round-window recorded potential of single-fibre discharge (unit response) in normal and noise-damaged cochleas. Hearing research 59:157-170.

Wang B (1979) The relation between the compound action potential and unit discharges of the auditory nerve. Doctoral dissertation, Dept of Electrical Engineering and Computer Sciences, MIT.

Weinberger MM, Kitzes LM, D.A. G (1970) Some characteristics of the auditory neurophonic. Experientia (Basle) 26:46-48.

Weiss TF, Rose C (1988a) A comparison of synchronization filters in different auditory receptor organs. Hear Res 33:175-180.

Weiss TF, Rose C (1988b) Stages of degradation of timing information in the cochlea: A comparison of hair-cell and nerve-fiber responses in the alligator lizard. Hear Res 33:167-174.

Zwislocki JJ, Feldman RS (1956) Just noticeable differences in dichotic phase. J Acoust Soc Am 28:860-864.

Chapter 5

Upper frequency limit of phase-locking in human and monkey

ABSTRACT

Phase-locking is a fundamental property of the peripheral auditory system. In single auditory nerve fibers, phase locking declines with frequency and becomes undetectable at an upper frequency limit, which differs between species. This limit is unknown in humans, and widely differing values are assumed in the interpretation of psychophysical results. Previous work in cat (chapter 4) showed that stimulus evoked mass potentials at the cochlear round window allow an assessment of the limit of neural phase-locking. We applied the same paradigm in macaque monkey and in human volunteers with normal hearing and recorded potentials from the cochlear promontory and niche of the round window through the middle ear. Neural phase-locking was quantified and the upper frequency limit was determined. The results suggest that the upper frequency limit of phase-locking in human is equal to or up to ~0.3 octaves lower than in monkey and cat.

INTRODUCTION

Neural phase-locking is the property of the auditory nerve (AN) fibers to lock to the sound waveform. Here we are concerned with phase-locking to the fine-structure of a sound waveform, which is most easily studied with pure tones. Classical studies showed that AN fibers tend to discharge at a preferred phase angle of a tone, and that this tendency decreases with increasing frequency (e.g., cat, Johnson, 1980; squirrel monkey, Rose et al., 1967). This behavior is traditionally quantified by the vector strength (VS) (Goldberg and Brown, 1969): VS decreases above a certain corner frequency and is not statistically significant beyond an upper frequency limit, which is species dependent (Weiss and Rose, 1988a). The upper frequency limit of phase-locking in humans is unknown. Current techniques to quantify phase-locking in single auditory nerve fibers cannot be applied in humans, because intracranial access is required with penetrating microelectrodes into the AN.

The relevance of phase-locking for perception has been debated for more than a century. It is clear that phase-locking is used in the binaural system of mammals with low-frequency hearing, where it contributes to spatial perception. Whether phase-locking subserves a more general role in auditory perception is a thorny issue. The upper limit of phase-locking provides a handle in this debate. Behavioral studies show that humans are sensitive to interaural temporal differences in fine-structure up to ~1.4 kHz (Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Schiano et al., 1986; Dunai and Hartmann, 2011).

Several authors argue that other (monaural) behavioral results indicate a temporal sensitivity to fine-structure up to frequencies as high as 8-10 kHz (Heinz et al., 2001; Moore, 2008), which is about an octave above the highest upper frequency limit found in single AN fibers in mammals. In contrast, we have argued (Joris and Verschooten, 2013) that the upper limit in humans is likely lower than that in the cat, by about an octave.

Obviously, physiological demonstration in humans with normal hearing of an absence of phase-locking above a certain frequency would rule out interpretations requiring coding of fine-structure above that frequency. Nevertheless, to our knowledge, an electrophysiological assessment of the upper limit of phase-locking in humans has not been attempted. As mentioned, recordings of single AN fibers in humans are presently not feasible. Mass potentials provide an alternative electrophysiological means of assessing the coding of fine-structure. Synchronous, phase-locked activity from AN fibers dominates the potentials measured on the surface of the AN in cat (Snyder and Schreiner, 1984) and human (Moller and Jho, 1989) and allow assessment of phase-locking limits (Chapter 4), but intracranial access is again required and precludes the study of healthy volunteers.

Here, we pursue a different approach which allows us to assess neural phase-locking in awake, normal hearing human volunteers and which combines two developments. First, studies in animals (Henry, 1995; He et al., 2012; Lichtenhan et al., 2013; our own studies: Chapter 4) show the presence of a phase-locked neural component in the electrical mass potential recorded at the cochlear round window (RW), suggesting this route as a means of studying neural phase-locking in humans. However, a problem arises because the alternating component of the mass potential at the RW is dominated by the cochlear microphonic (CM), which is generated by the hair cells (Dallos, 1973) and which extends to much higher frequencies than phase-locking in the AN (Snyder and Schreiner, 1984). We developed a method (Chapter 4) to disentangle these two components, thus enabling measurement of neural phase-locking at the RW. Our tests in a previous study in cat demonstrated the validity of this method and showed a good correspondence between phase-locking limits reported in single AN fibers and the limit found with our method. Second, we adapted a recording method used in clinical practice towards use in volunteers with normal hearing. Although it sees only limited use today, a minimal invasive approach enables the recording of cochlear potentials in patients for diagnostic purposes, using a transtympanic needle placed on the cochlear promontory or niche of the RW (Eggermont, 1976a; Harrison et al., 1981). We adapted this method (Chapter 3) to record stimulus evoked potentials in awake normal-hearing volunteers over long time periods (up to ~4 hours) using a closed, calibrated acoustic system. The same paradigm and extraction method previously developed in cat (Chapter 4),

was then applied to quantify the auditory neurophonic and to determine the detectable upper frequency limit of phase-locking.

Finally, experiments were conducted in macaque monkey using a similar transtympanic approach. Phase-locking in single AN fibers of macaques was previously studied in our laboratory (Michelet et al., 2011): the corner frequency is somewhat lower than in cat with a maximal difference of ~0.3 octave. As we will argue in DISCUSSION, the combined availability of both single-fiber and mass-potential data in monkey provides significantly strengthens our interpretation of the human data.

METHODS

All procedures, for both animal and human experiments, were approved by the Medical Ethics Committee and Animal Ethics Committee of the University of Leuven and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Subjects

Humans. The volunteers who participated in this study were between 20 and 35 years old and recruited from the local student population; all gave written informed consent. Measurements were obtained from 11 subjects, but were retained for only 6 subjects (5 female, 1 male). In the 5 other subjects, either the signal-to-noise ratio (SNR) was very poor or the maximum measured probe frequency was too far below the upper frequency limit. More information about the subjects, experimental setup and transtympanic procedure can be found in METHODS of (Chapter 3). The recordings reported here were interleaved with measurements of frequency tuning (Chapter 3).

Monkeys. Recordings were obtained from one ear in three monkeys (Macaca mulatta) which were also used in chronic visual experiments (an 8.9 kg adult male, a 4.8 kg juvenile male; and a 6.3 kg juvenile female). Prior to the experiments, dissections on formaline-preserved temporal bones were performed to study the desired trajectory and placement of the needle electrode. The recording procedure in monkey was similar to that in human, with the main difference being the presence of anesthesia. Induction of general anesthesia was with a mixture of ketamine (3 mg/kg) and medetomidine (0.050 mg/kg), i.m. The same mixture was administered i.v. for maintenance, through a venous cannula inserted for administration of lactated Ringer's solution. The tympanic membrane was checked otoscopically. A custom ear

mold was made. The needle electrode was placed under guidance of a surgical microscope and was fixed relative to the head with a mechanical manipulator. The duration of the total experimental session including the placing of the electrode varied between 4 and 6 hours. After the experiments, atipamezole (0.2 mg/kg) was administered to reverse the sedative effect of medetomidine; after awakening of the animal, it was observed for a certain period of time.

Stimulus generation

Stimuli were generated with custom software and a digital sound system (Tucker-Davis Technologies, system 2, sample rate: 125 kHz/channel) consisting of a digital-to-analog converter (PD1), a digitally controlled analog attenuator (PA5), a headphone driver (HB7) and an electromagnetically shielded earphone-speaker (Etymotic Research, ER1 or ER2, 20 Hz - 16 kHz). The transducer was connected with plastic tubing to the custom ear mold. The complete acoustic system was calibrated *in situ*, with a probe-microphone (Etymotic Research, ER-7C) close to the tympanic membrane.

Signal recording

In human and monkey, the same recording equipment was used as in "Electrophysiological recording and analysis" of the METHODS in chapter 3. The only difference with the latter study was a broader bandwidth for the analog band-pass filter (0.03 - 30 kHz).

Stimulus Paradigm

The method used to extract the neurophonic from the potentials in human, and monkey and to disambiguate it from CM, is based on forward masking and stimulus polarity reversals. The method is almost identical to the one previously developed in cat, including the Gabor transform to determine the magnitude (see METHODS in chapter 4). The paradigm applied was also the same, but some parameters were modified in order to optimize the measurement time. First, the length of the probe was reduced and fixed to 50 ms and the subsequent probe-masker intervals were reduced to 10 ms. Second, to cope with the lower SNR (relative to that in cat, the number of averages (n >= 200) and the stimulus levels were increased. The probe levels in the different experiment were higher than in cat; 65, 70 or 75 dB SPL versus 50 or 55 dB SPL in cat. For efficient masking, the masker level was fixed at 10 dB above the probe level. Fig. 1 illustrates the paradigm (Fig. 1a,c) and average responses (Fig. 1b,d) from 1 human subject. The upper (Fig. 1a) and lower (Fig. 1c) half stimulus representations are

basically equal, but with opposite stimulus polarity. One half representation consists of 3 segments. The first segment A contains only a probe; the second segment B contains a masker followed by a probe; and the last segment C contains only the masker. The probes (50 ms) and maskers (83.71 ms) were pure tones equal in frequency (f_P). The masker-probe interval was 1 ms, and the interval between different stimuli was at least 10 ms. Segment C had a stimulus-free period of ~80 ms, which was required for the recovery of masking and was also used for the determination of the noise floor in the individual results. To reduce spectral splatter the probe and masker were gated with a 1 ms raised-cosine.

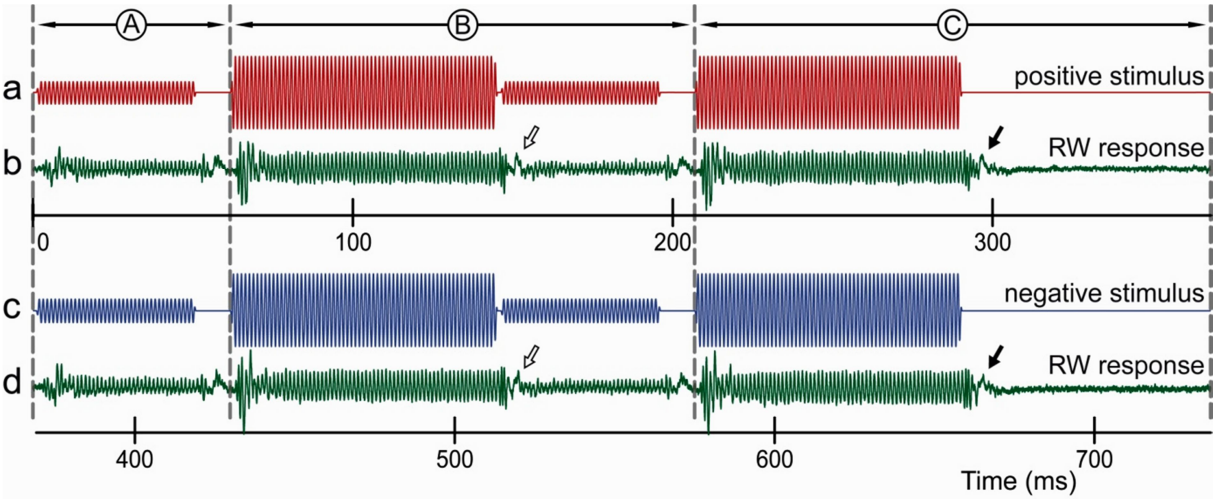


Figure 1. An illustration of the paradigm (a,c) using the 3 averaged segments of a real stimulus (A,B,C), with corresponding averaged responses at the promontory in human (b,d). Section A contains the probe only, secton B contains the probe preceded by a masker, and section C contains the masker only. The probe and masker are pure tones of 800 Hz; the phase in c (blue) is opposite to that in a (red). The arrows indicate the offset response of the masker in B and C.

Analysis

The analysis performed on the responses of human and monkey was the same as in the previous study in cat (METHODS, chapter 4), in which the responses to different segments were combined to obtain the required signals. We summarize this with the aid of Fig. 1. The first operation always performed, was the removal of the masker's trailing off-set response (filled arrows in segment C of Fig. 1b,d) from the forward masked probe response (open arrows in segment B Fig. 1b,d). We refer to the corrected response simply as the *masked response*. The unmasked probe response (segment A in Fig. 1b,d) is referred to as the *probe response*.

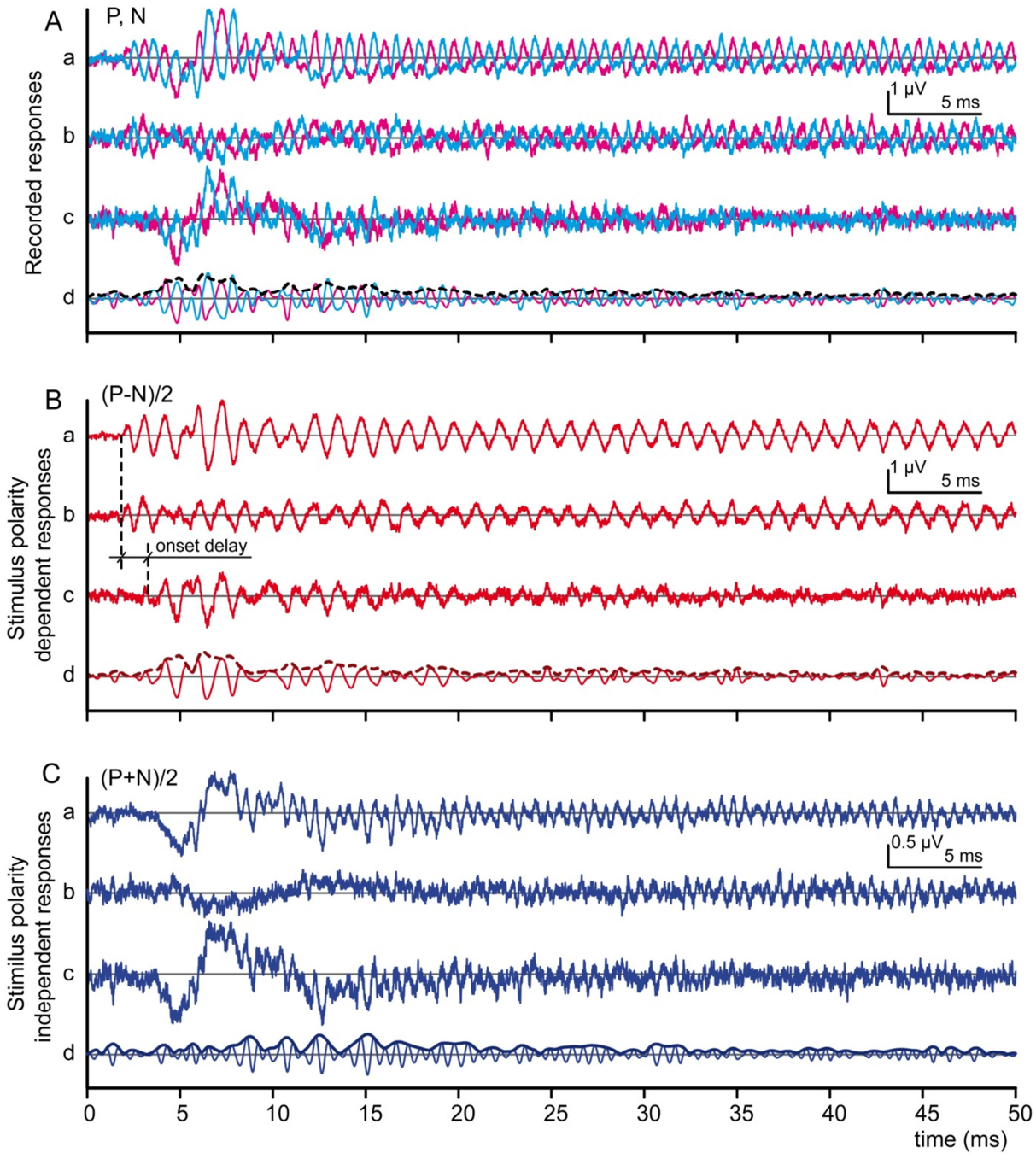


Figure 2. Example of averaged stimulus evoked responses recorded in human. Aa) A pair of raw probe responses for opposite stimulus polarity. Ab) Same as Aa, but with a preceding tonal masker. Ac) Difference between signal Aa and Ab: the adapted component. Ad) Same as Ac, but with the CAP removed by high-pass filtering. Ba-d) Difference of the responses shown in A. Ca-d) Sum of the responses shown in Aa-d. Parameters: probe level = 75 dB SPL, masker level = 85 dB SPL, frequency = 800 Hz. The final traces in each panel (Ad,Bd,Cd) are denoised for illustrative purposes, and their envelope is shown as the dashed line.

Fig. 2 illustrates the combinations of the responses from the different segments (A,B,C) in Fig. 1. The first two pairs of responses are the basic alternated (positive, P, magenta; negative, N, cyan) probe responses (Fig. 2Aa) and masked responses (Fig. 2Ab). From these pairs of basic responses a variety of signals and magnitude envelopes are derived. The first signal (pair) is the *adapted component* (Fig. 2Ac), which is the masked component in the potential and is obtained by subtracting the masked response (Fig. 2Ab) from the probe response (Fig. 2Aa). The high-pass (cut-off at $\frac{1}{2}f_P$) filtered and de-noised (only for the illustration) versions of the adapted components and its mean envelope (dashed line) are shown in Fig. 2Ad; we refer to these signals as the decaying neurophonic (dNP). Next, the subtraction of the halved pairs of responses [(P-N)/2] in Fig. 2Aa,b,c results in the *stimulus polarity dependent responses* (Fig. 2Ba,b,c) and the summation of these [(P+N)/2] results in *stimulus polarity independent responses* (Fig. 2Ca,b,c). Similar to Fig. 2Ac, the filtered and de-noised versions with their envelopes of the signals in Bc and Cc are represented by Bd and Cd; the low-pass cut-off frequencies of the filters were $\frac{1}{2}f_P$ (Bd) and f_P (Cd). The maximum magnitude over the time course of the dNP was extracted with the same method (Gabor transform) as described in chapter 4, with exactly the same settings. The maximum amplitude of the 1st harmonic was derived with the latter method from the stimulus polarity dependent dNP (cf. Fig. 2Bd), the maximum amplitude of the 2nd harmonic from the stimulus polarity independent dNP (cf. Fig. 2Cd) and the maximum amplitude of the total contribution (combined powers of the 1st, 2nd, 3rd, and 4th harmonic) from the mean power spectrum of the pairs from the dNP (cf. Fig. 2Ad). The corresponding noise floors in these signals were calculated, as described in chapter 4, from the background noise in the artifact free part of segment C (cf. Fig. 1b,d).

RESULTS

We present data obtained in 3 monkeys and in 6 human volunteers. Measures for the total phase-locking and its 1st and 2nd harmonic were obtained from the adapted responses (cf. 2Ad,Bd,Cd) using the Gabor method. The neurophonic was obtained for a range of frequencies in each subject; frequencies were presented in an arbitrary sequence.

Humans

Time course. Fig. 2 illustrates the typical unmasked (Fig. 2Aa) and masked pairs (Fig. 2Ab) of probe responses for a frequency of 800 Hz, with their derived signals (Fig. 2Ac,d; Fig. 2B; Fig. 2C). One of the most notable differences with respect to the recordings in cat (Fig. 5 in chapter 4) is that the amplitudes in human are generally much smaller (by ~20 dB) and that the time course of the dNP has some additional modulation. The probe responses (Fig. 2Aa) in human show all basic features expected (Figure 5 in Chapter 4), including a stimulus polarity dependent and independent alternating component (Fig. 2Ba,Ca) and a stimulus polarity independent CAP (Fig. 2Ca). These features are reduced in amplitude following the forward masker (Fig. 2Ab, Fig. 2Bb and Fig. 2Cb). As we previously found at the RW of the cat, masking is incomplete in the stimulus polarity dependent component (Fig. 2Ab,Bb) due to the presence of a non-maskable component, the CM. The adapted component (Fig. 2Ac,Bc,Cc), or better, its CAP-filtered version which we call the decaying neurophonic (dNP) (Fig. 2Ad,Bd,Cd), reveals a significant neural contribution. This contribution is large and is of the same order of magnitude as the CM (see below). As expected, an onset lag (here ~1.5 ms) is present between the (un)masked response (Fig. 2Aa,b) and the adapted component (Fig. 2Ac), which indicates that the CM in the masked response is effectively cancelled and that the adapted component mainly consists of neural contributions. In other words, the unmaskable component in the response before neural onset (4 ms) is the CM.

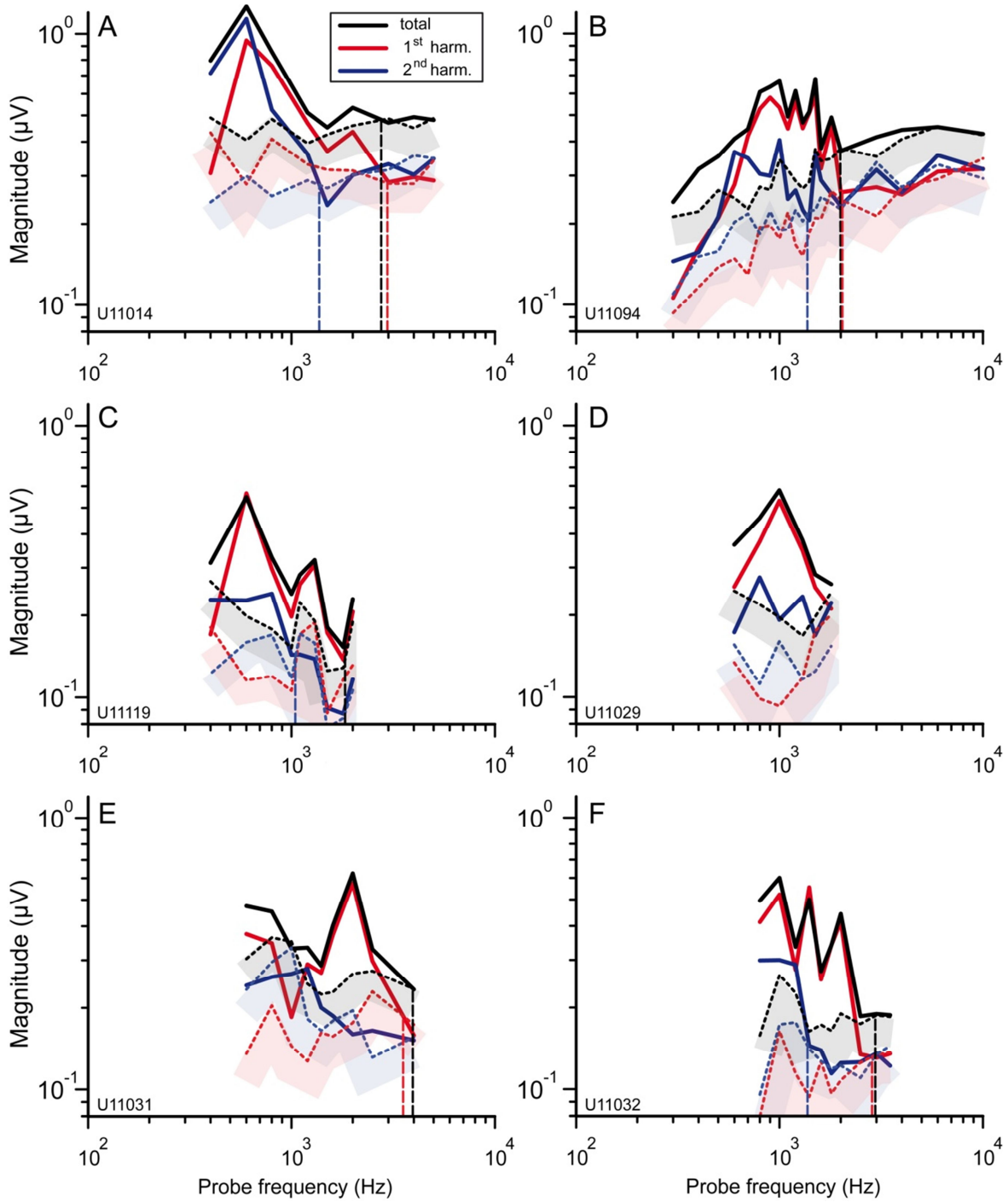


Figure 3. Neurophonic as a function of probe frequency obtained near the RW in 6 humans. The neurophonic was obtained as the maximum magnitude of the time course of the decaying neurophonic using the Gabor method. The curve of the total (black) harmonic contribution was obtained from the first 4 harmonic components; the others from the 1st (red) and 2nd (blue) harmonic. The noise floor of the neurophonic is indicated by the dotted lines. C is from a male subject, all other panels are from female subjects.

Frequency dependence. Fig. 3 presents the neurophonic for a range of probe frequencies obtained in 6 humans. The neurophonic was obtained as the maximum magnitude of the time course of the decaying neurophonic (dNP) using the Gabor method. The magnitudes of the total harmonic contribution (black; summation of the powers of the first 4 harmonics) were obtained from the decaying pairs (cf. Fig. 2Ad); these of the 1st harmonic (red) were obtained from the stimulus polarity dependent response (cf. Fig. 2Bd) and these of the 2nd harmonic (blue) from the stimulus polarity independent response (cf. Fig. 2Cd). All curves show the same trend: the magnitude increases with frequency until it reaches an absolute maximum, thereafter it decreases until it is bounded by its noise floor. Above a probe frequency of ~800 Hz the total neurophonic is dominated by the 1st harmonic; below ~800 Hz, both the 1st and 2nd harmonic show a substantial contribution. The detectable upper frequency limit is indicated by the vertical dashed lines and is the frequency where the neurophonic meets its noise floor. The frequency limits vary between subjects. The limit of the total contribution (black) and 1st harmonic (red) are between 2 and 4 kHz and these of the 2st harmonic (blue) between 1 and 1.5 kHz. Some limits could not be determined due to a lack of data (Fig. 3D) or because of ambiguity (Fig. 3E, 2nd harmonic). A general limit for this human dataset is obtained below (section "Population analysis")**.** Note that at the promontory in human the signal to noise floor ratio (SNfR) is rather low with a maximum value of roughly 10 dB, which is much lower than in cat, where it is ~30 dB. The noise floor of the total contribution (upper black dashed line) is considerable higher than that of 1st and 2nd harmonic; this is due to the larger bandwidth and therefore spectral noise power.

Monkeys

The amplitude of the neurophonic in monkey was in general somewhat larger than in human, but the noise floor was comparable. The time courses of the probe responses, masked probe responses and adapted components in monkey were very similar to those of human. Again, the same features (phase-locked components, CAP, decaying course and onset delay) were present. An example of the typical time course of the decaying neurophonic (cf. Fig. 2Ad) is shown in Fig. 4.

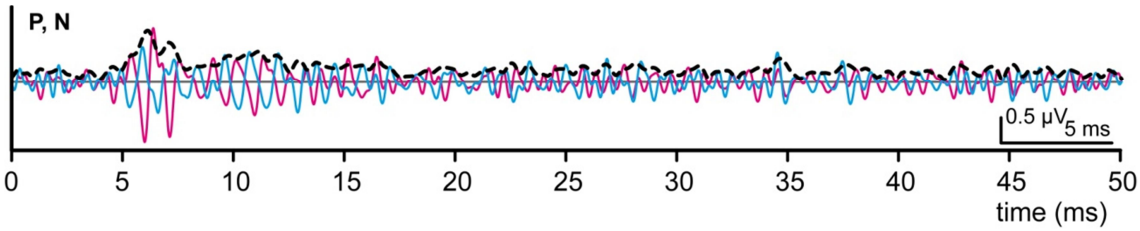


Figure 4. Example of a typical time course of the decaying neurophonic (cf. Fig. 2Ad) in macaque monkey. The frequency was 800 Hz.

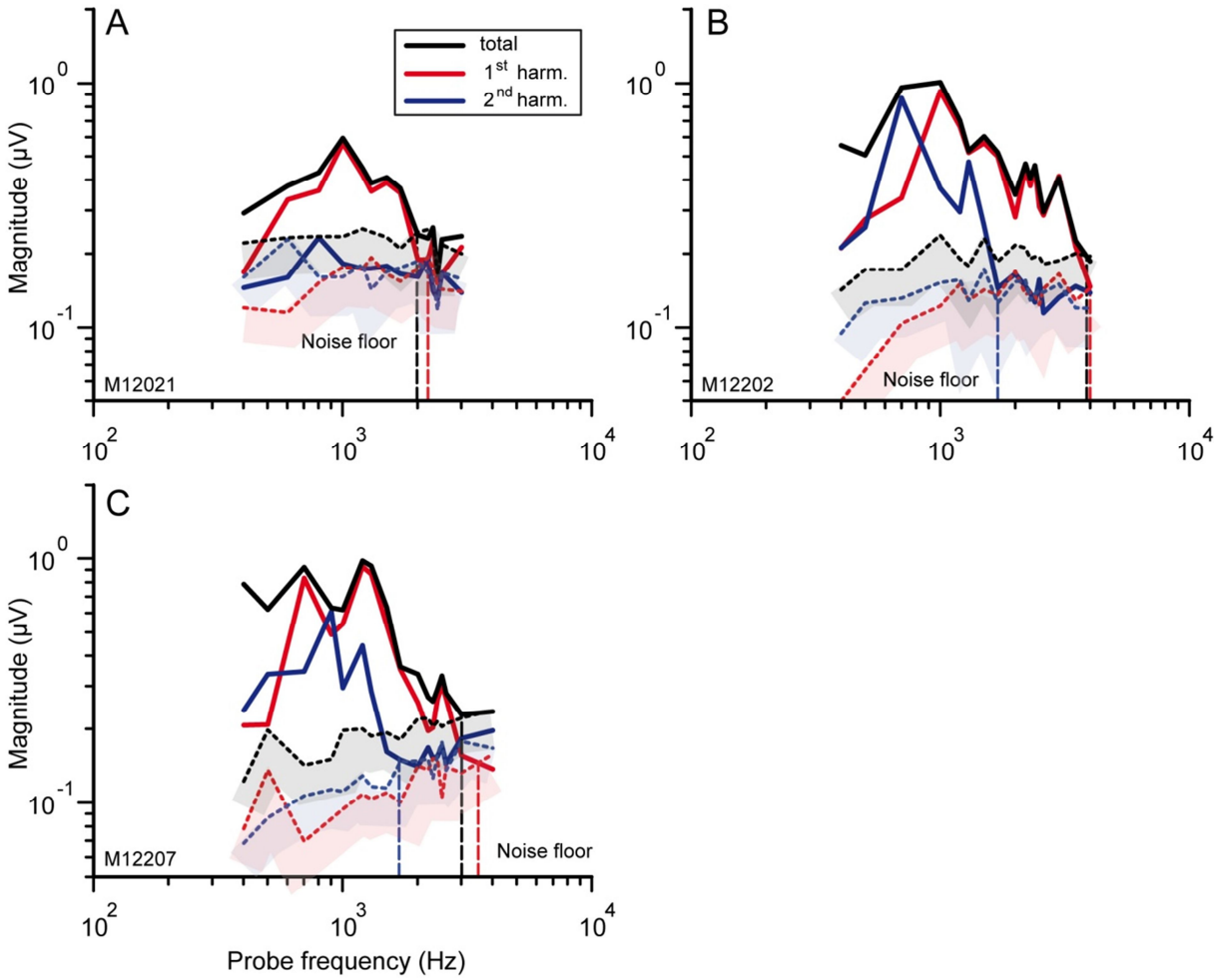


Figure 5. Same as Fig. 3 but for 3 macaque monkeys. A was an adult male Rhesus monkey. B and C were male and female juveniles, respectively.

Fig. 5 shows the trends of the neurophonic as a function of probe frequency for 3 monkeys. The magnitudes of the neurophonic and noise floors were obtained in the same way as in human. The trend lines show a band-pass characteristic similar to that in human, with a center frequency near 1 kHz. Above this frequency the 1st harmonic is dominant; below it, the

2nd harmonic is also prominent. The maximum detectable upper frequency limit of phase-locking (vertical dashed lines) is between 2 and 4 kHz for the total and 1st harmonic, and between 1.6 and 2 kHz for the 2nd harmonic.

Population analysis

In this section we combine the results within each data set (human and monkey) to determine an overall upper frequency limit of phase-locking. Fig. 6 shows the single data points and trends lines (MATLAB, *LOESS*, span: 0.55) of the total (black), 1st (red) and 2nd (blue) harmonic of the maximum amplitudes of the dNP and its noise floor, for each data set. The trend lines for human (panel A) and monkey (panel B) reflect the underlying typical band-pass characteristic of the individual experiments. The center frequency for human and monkey is 0.7 and 1 kHz, respectively. Remark that only two experiments contributed data above 4 kHz, which biases this part of the trend lines towards those experiments (Fig. 3A and B).

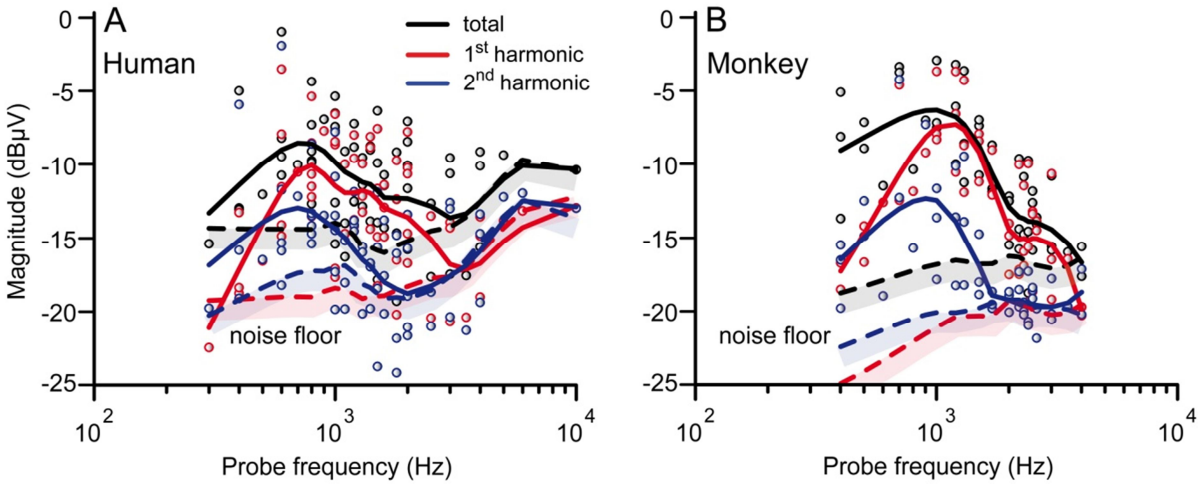


Figure 6. Population trendlines for the neurophonic (solid) and noise floors (dashed and shaded). A) For the 6 experiments in human (Fig. 3). B) For the 3 experiments in monkey (Fig. 5). Single data points are indicated by colored dots; the total contribution (first 4 harmonics) in black; 1st harmonic in red and 2nd harmonic in blue. The trend lines were obtained with a *LOESS* fit (MATLAB, span: 0.55).

Because of the low SNfR in human (max. 8 dB for the 1st harmonic) and to a lesser extent in monkey (max. 15 dB for 1st harmonic), the noise floor has a clear effect on the course of the trendlines. This is clearly the case at the highest frequencies where the neurophonic is overtaken by the noise floor, i.e. each solid trendline converges with its dashed corresponding noise floor. However, also below the point of convergence, noise contributes significantly to the magnitude of all trendlines due to the mere size of the SNfR. In cat, which

has a much higher SNfR, the trendlines of the total and 1st harmonic contribution essentially superimpose at mid and high frequencies, where the 2nd harmonic is very small (chapter 4). The general tendency of trendlines in human and monkey shows surprisingly similarities to that in cat (see DISCUSSION), including a small 2nd harmonic at mid and high frequencies, but the total harmonic contribution (black curve) is clearly raised relative to the 1st harmonic at mid and high frequencies (Fig. 6). Clearly, noise contributes significantly at all frequencies and has to be taken in to account to properly interpret the trendlines.

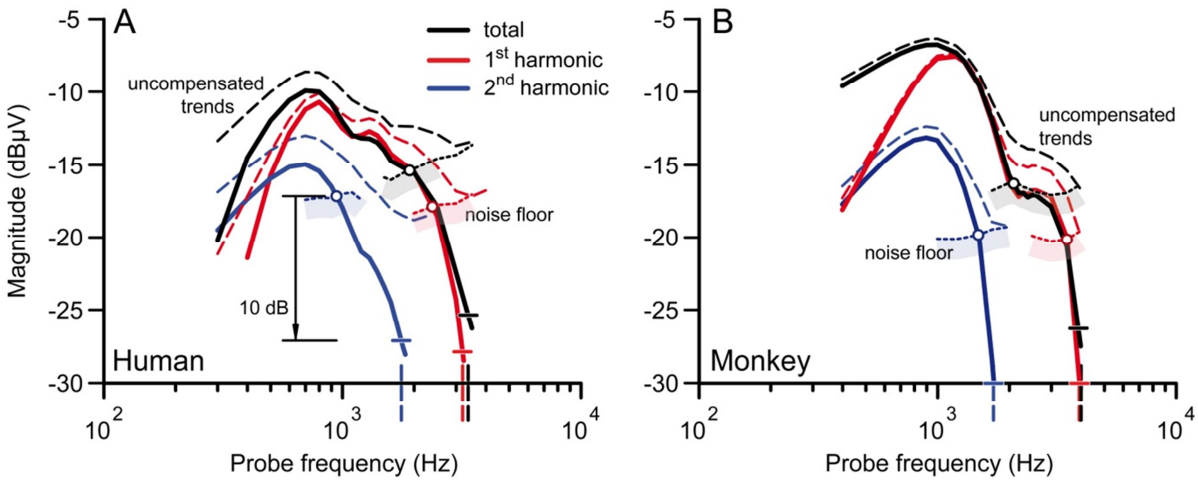


Figure 7. Compensated trendlines of the neurophonic (solid) corrected for the influence of the noise floor (dotted and shaded) for A) Human and B) Monkey. The uncompensated trendlines of the neurophonic (dashed lines) in Fig. 6 are included for comparison. The compensated lines were obtained by subtracting the power of the noise floor. The total contribution (first 4 harmonics) is indicated in black; 1st harmonic in red; and 2nd harmonic in blue. White dots: intersection of compensated trendlines with noise floor. Short horizontal and vertical lines at bottom of compensated trendlines: estimated upper frequency limits.

Fig. 7 shows compensated trendlines (solid) together with the original trends (dashed lines) and the noise floors (dotted lines). The compensation involved subtraction of the power of the noise floor from the power of the neurophonic under the assumption of absence of correlation. The compensation has a clear correcting influence on the course of the curves, especially in human and for those frequencies were the SNfR is small. For the frequencies where the contribution of the 2nd harmonic is small (>1 kHz), the compensated trend lines of the total (black) and the 1st harmonic contribution (red) are now almost exactly the same. Given the large differences in the uncompensated trendlines (Fig. 6), this result is not trivial and attests to the validity of our approach.

Inherently, due to inaccuracies in measurement of the neurophonic and its noise floors, the reconstruction of the course of the compensated trendlines is only reliable for those frequencies where the compensated trend line is not too far below the noise floor. By definition, the compensated curves go to minus infinity at the upper detectable frequency limit where the uncompensated trendlines first touch their noise floors. At this frequency and just before it, the course of the compensated curve is unreliable. As in our previous study (chapter 4), we only retained the part extending 10 dB below the intersection with the noise floor (indicated by the white dots and dashes in Fig. 7); the frequency at the end of the curve is here defined as the upper frequency limit of phase-locking (indicated by the horizontal dashes on the compensated trend lines). Using the latter definition we obtained the following upper frequency limits: for the 2nd harmonic 1.8 ±0.1 (SEM) kHz in human and 1.7 ±0.2 kHz in monkey; and for the 1st harmonic (and total contribution) 4.0 ±0.6 kHz in monkey and 3.3 ±0.3 kHz in human, which is 0.28 octave lower than monkey. Note that the SEM is calculated as the product of the estimated upper frequency limit of the population with the relative SEM from the upper frequency limits of the original neurophonic curves..

DISCUSSION

The extent of auditory neural phase-locking in human is unknown and is debated (Heinz et al., 2001; Moore, 2008; Joris and Verschooten, 2013). The aim of the present study was to electrophysiologically assess peripheral phase-locking in human. A first important result of our study is that it is possible to identify a phase-locked neural component in recordings *via* the human middle ear. The existence of such a component at the RW of animals had been shown in previous studies (Henry, 1995; He et al., 2012; Lichtenhan et al., 2013) and was fully characterized in our previous study (Chapter 4). We adapted a clinical approach for measurement of cochlear potentials in humans so that we could sample potentials over an extended period of time, with full acoustic control, in volunteers with normal hearing. We also applied the same technique and analysis to the macaque monkey, for which single AN fiber recordings are available in our laboratory, so as to strengthen the interpretation of the results in human and to tighten the correlations between neurophonic and single fiber data.

In many respects, the neural potentials in human and monkey behaved similar to those previously recorded at the RW in cat (Chapter 4). In all species, the potentials show forward

masking, adaptation, and an onset delay. Also, as a function of frequency, the overall pattern of results has basic similarities across species (compare Fig. 7 with Fig. 22 in Chapter 4). The maximum amplitude of the dNP shows strong dependence on probe frequency, with a clear band-pass characteristic centered, roughly, on 800 Hz. Above this frequency the dNP is dominated by the 1st harmonic; below it the 2nd harmonic gains in prominence and even dominates at the lowest frequencies tested (Fig. 6 and Fig. 7). We previous extensively discussed the evidence that the measure we adopted, the dNP, is generated by the phase-locked activity of single AN fibers (Chapter 4). Compared to the cat, the data obtained in human and monkey are more limited in volume and SNfR, and less variables were explored (e.g. SPL, effect of TTX), but the data available strongly suggest that the dNP measured in human and monkey is also uniquely generated by phase-locked neural activity.

Comparison of the upper frequency limit

An important goal of our study was to determine the upper frequency limit of phase-locking in human. In a previous study in cat (Chapter 4) we found good agreement between the detectable frequency limit of phase-locking obtained from RW-potentials i.e., 4.7 kHz for the 1st harmonic, and that from the VS of single nerve-fibers in cat i.e. ~5 kHz (Johnson, 1980; Joris et al., 1994b). In the present study we obtained a lower detectable frequency limit for human (3.3 kHz) with respect to cat but also with respect to monkey (4 kHz). Based on these population results one may conclude that the upper *detectable* frequency limit of human is 0.28 octave lower than monkey and 0.51 octave lower than cat. To what extent do these limits reflect real differences in upper frequency limit of neural phase-locking? This raises the fundamental issue of the relative positions of the trendlines.

We start by showing (Fig. 8A) the absolute magnitude of the compensated trendlines for the fundamental (1st harmonic) of the neurophonic for the 3 species. The trendlines of human and monkey are taken from Fig. 7, the cat data are from Fig. 26 in chapter 4. There is clearly a large difference in magnitude between cat (blue) *versus* human (red) and monkey (green), which are offset by ~30 dB. At least some of this difference must be attributable to the site of recording, which was on the RW in cat but more remote in human and monkey. Thus, a comparison purely in terms of "raw" magnitude favors recordings in the cat, where the RW is easily accessed, and may bias the frequency limit upward in this species, relative to human and monkey. We know that this must be the case for the comparison cat-monkey, since the upper limit in single fiber recordings in monkey (Rose et al., 1967) and particularly in macaque monkey (Michelet et al., 2011) is similar.

Note that the noise floors at their intersection with the compensated trend lines (dashed lines, Fig. 8A, compare with white dots in Fig. 7), are rather similar in magnitude in the 3 species, and are not ranked in the same way as the overall compensated trendlines. This suggests that these noise floors near the upper limit are not dominated by neural noise from the cochlea, which peaks near 1 kHz (Dolan et al., 1990; Patuzzi et al., 2004). Thus, a first reasonably comparison of the curves is by using the noise floor as a reference to position the different curves (Fig. 8B). This is essentially the procedure illustrated in Fig. 7 resulting in the values quoted above.

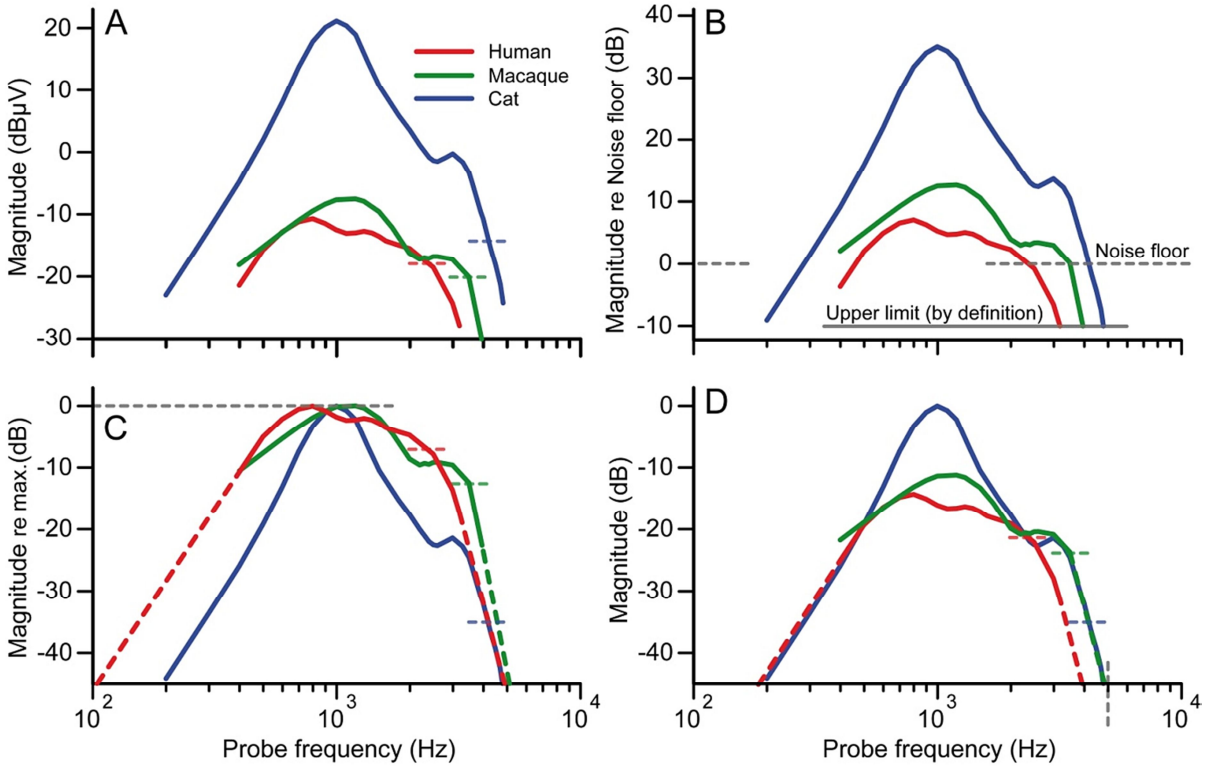


Figure 8. Comparison amongst the neurophonic of human, monkey and cat for different relative positions. A) For aligned noise floors. B) For normalized maximum magnitudes. C) For corrected high-frequency slope of macaque, while maintaining the relation between macaque and human. The slopes of monkey and human are extended (thick dashed lines). The intersections of the neurophonic with their noise floors are indicated by the horizontal dashed lines. The upper limit in B indicated by the gray line is by definition 10 dB lower than the noise floor.

Average maximal VS values in monkey and cat are very similar for frequencies up to ~1 kHz (Michelet et al., 2011). Under the reasonable assumption that this similarity would extend to human, one could argue that the compensated trend lines should be compared normalized to their maximum, as shown in Fig. 8C. Where necessary, the slopes are extended (dashed lines) to the abscissa (here 10 dB below the noise floor of cat). View this way, both

the upper limit and the slope near that limit are similar between species, but the lower SNfR in human and monkey forces the neurophonic to merge earlier with its noise floor resulting in an underestimation of the upper limit. However, the comparison in Fig. 8C is also faulty. First, we know from the single-fiber data in cat and macaque monkey that – if anything – the upper limit is lower rather than higher in monkey re. cat. Second, the VS single-unit data are essentially flat for frequencies <1 kHz, while the trendlines all show a high-pass slope. Clearly, other factors than VS influence the shape of the trendlines.

In the final panel (Fig. 8D), we made the assumption that all species show the same high frequency slope. This is already the case for the compensated trendlines of cat and human in Fig. 8C; the slope is somewhat steeper in monkey, and is adjusted to equal that of cat and human in Fig. 8D. Moreover, for simplicity, we align the high frequency slopes of monkey and cat, in essence assuming equality at those frequencies in the two species (even though the single unit data show that VS values in the monkey are slightly lower than in cat). We maintain the relationship between monkey and human (as in Fig. 8A). Interestingly, these manipulations result in identical high-pass slopes in human and cat. In this case the upper frequency limit in human is estimated at 4 kHz i.e. 0.31 octave lower than in cat and monkey.

That monkey and human are more similar to each other than to the cat (in neurophonic amplitude, noise floor, and of course in anatomy), combined with the observation that (the high-frequency slopes) single-fiber VS is ~0.3 octave lower in macaque monkey than in cat, makes us surmise that the real relationship between the distributions is closer to Fig. 8B,D than to the extreme in Fig. 8C. A conservative conclusion is that our measurements provide no support for the hypothesis that neural phase-locking in humans extends higher than in cat. We believe our data argue that the limit in humans is actually lower, though not fully by an octave as we proposed earlier (Joris and Verschooten, 2013).

Not only neural phase-locking *per se* (i.e. the property of synchronization of spikes to the stimulus waveform, as expressed by VS), but many other factors contribute to the magnitude of the neurophonic and its trend with frequency. We have already mentioned anatomical factors such as site of recording and anatomy of the cochlea and surrounding structures, which may affect the "unit contribution" i.e. how a single action potential contributes to the electrical signal recorded. Also, the cochlea of monkey and human is biased towards lower frequencies than in cat and their frequency tuning is sharper (Shera et al., 2002; Joris et al., 2011, Chapter 3), so that the spatiotemporal pattern of phase-locked fibers along the basilar membrane likely differs with the cat. Other factors that may differ between these species and affect the neurophonic results are neural adaptation, hearing threshold at different

frequencies, the difference in SPLs used, proximity of the cochlear nucleus, etc. Such factors could cause differences in the neurophonic trendlines, even if the dependence of VS on frequency would be identical in human and cat AN fibers.

In conclusion, this study shows that it is possible to extract neural phase-locking from cochlear potentials at the cochlear promontory or in the niche of the round window in volunteers. The upper frequency limit of phase-locking obtained in the present study was 4.0 kHz for macaque monkey and 3.3 kHz for human, which is 0.28 octaves lower than monkey and 0.51 octaves lower than cat (4.7 kHz). The lower limit of the neurophonic of macaque monkey (4 kHz) with respect to its maximum VS (5 kHz) suggests that the neurophonic in monkey, and therefore also in human, slightly underestimates the actual phase-locking limit due to the low SNfR of the neurophonic. Thus, the upper frequency limit of phase-locking in human might be closer to 4 kHz than to 3.3 kHz.

A serendipitous benefit of our procedure of recording in humans is the absence of general anesthetics, while all AN single-fiber studies are performed in anesthetized animals. At a general level, the ability to access neural phase-locking demonstrated here opens up the possibility to study this property in human conditions where a disturbance of fine temporal coding has been proposed to cause hearing problems, such as in neuropathy, tinnitus, and multiple sclerosis (Moller, 1984; Levine et al., 1993; Zeng et al., 2005).

CONCLUSIONS

The same paradigm developed in Chapter 4 was applied in monkey (macaque) and in human volunteers with normal hearing. Potentials were recorded from the cochlear promontory and niche of the round window. Neural phase-locking was extracted and quantified. We found that the temporal and spectral properties regarding phase-locking in human, monkey and in cat were very similar. The most notable difference in human and monkey compared to cat was a much smaller magnitude (-30 dB) and signal-to-noise(floor) ratio for the neurophonic. The upper detectable frequency limits of neural phase-locking obtained from stimulus evoked potentials was lower in human (3.3 kHz) than in macaque (4.0 kHz). Comparison between the latter limit in monkey and that of single AN fibers suggests that the limit for AN-fiber in human is somewhere between 3.3 and 4.0 kHz and therefore lower than in cat. A conservative conclusion is that our measurements provide no support for the hypothesis that neural phase-locking in humans extends higher than in cat. We believe our data argue that the limit in humans is actually lower, though not fully by an octave as we proposed earlier (Joris and Verschooten, 2013).

REFERENCES

Dallos P (1973) The auditory periphery. New York: Academic Press.

Dolan DF, Nuttall AL, Avinash G (1990) Asynchronous neural activity recorded from the round window. The Journal of the Acoustical Society of America 87:2621-2627.

Dunai L, Hartmann WM (2011) Frequency dependence of the interaural time difference thresholds in human listeners. J Acoust Soc Am 129:2485–2485.

Eggermont JJ (1976a) Electrophysiological study of the normal and pathological human cochlea. II. Neural responses. Rev Laryngol Otol Rhinol (Bord) 97 Suppl:497-506.

Goldberg JM, Brown PB (1969) Response of Binaural Neurons of Dog Superior Olivary Complex to Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. J Neurophysiol 22:613-636.

Harrison RV, Aran JM, Erre JP (1981) AP tuning curves from normal and pathological human and guinea pig cochleas. J Acoust Soc Am 69:1374-1385.

He W, Porsov E, Kemp D, Nuttall AL, Ren T (2012) The group delay and suppression pattern of the cochlear microphonic potential recorded at the round window. PLoS One 7:e34356.

Heinz MG, Colburn HS, Carney LHC (2001) Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. Neural Comput 13:2273-2316.

Henry KR (1995) Auditory nerve neurophonic recorded from the round window of the Mongolian gerbil. Hear Res 90:176-184.

Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J Acoust Soc Am 68:1115-1122.

Joris PX, Verschooten E (2013) On the limit of neural phase-locking to fine-structure in humans. Basic Aspects of Hearing: Physiology and Perception Editors: BCJ Moore, RD Patterson, I Winter, RP Carlyon, and HE Gockel (Springer, NY), in press.

Joris PX, Carney LH, Smith PH, Yin TC (1994b) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. J Neurophysiol 71:1022-1036.

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proceedings of the National Academy of Sciences of the United States of America 108:17516-17520.

Klumpp R, Eady H (1956) Some measurements of interaural time differences thresholds. J Acoust Soc Am 28:859-864.

Levine RA, Gardner JC, Fullerton BC, Stufflebeam SM, Carlisle EW, Furst M, Rosen BR, Kiang NY (1993) Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials. Hear Res 68:73-88.

Lichtenhan JT, Cooper NP, Guinan JJ, Jr. (2013) A new auditory threshold estimation technique for low frequencies: proof of concept. Ear Hear 34:42-51.

Michelet P, Mc Laughlin M, van der Heijden M, Joris P (2011) Synchronization to pure and amplitude-modulated tones in the auditory nerve of macaque monkey. Assoc Res Otolaryngol Abstract 34:223.

Moller AR (1984) Pathophysiology of tinnitus. The Annals of otology, rhinology, and laryngology 93:39-44.

Moller AR, Jho HD (1989) Response from the exposed intracranial human auditory nerve to low-frequency tones: basic characteristics. Hear Res 38:163-175.

Moore BC (2008) The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. J Assoc Res Otolaryngol 9:399-406.

Patuzzi RB, Brown DJ, McMahon CM, Halliday AF (2004) Determinants of the spectrum of the neural electrical activity at the round window: transmitter release and neural depolarisation. Hear Res 190:87-108.

Rose JE, Brugge JF, Anderson DJ, Hind JE (1967) Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J Neurophysiol 30:769-793.

Schiano JL, Trahiotis C, Bernstein LR (1986) Lateralization of low-frequency tones and narrow bands of noise. J Acoust Soc Am 79:1563-1570.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proceedings of the National Academy of Sciences of the United States of America 99:3318-3323.

Snyder RL, Schreiner CE (1984) The auditory neurophonic: basic properties. Hearing research 15:261-280.

Weiss TF, Rose C (1988a) A comparison of synchronization filters in different auditory receptor organs. Hear Res 33:175-180.

Zeng F-G, Kong Y-Y, Michalewski HJ, Starr A (2005) Perceptual consequences of disrupted auditory nerve activity. J Neurophysiol 93:3050-3063.

Zwislocki JJ, Feldman RS (1956) Just noticeable differences in dichotic phase. J Acoust Soc Am 28:860-864.

Chapter 6: General Discussion

Decomposition of sound into bands of its constituent frequencies, and coding of the temporal waveform of these bands of frequencies, are two fundamental properties of the peripheral auditory system. All treatments of how sound is coded and analyzed by the central nervous system ultimately refer to these two properties. It is an important hiatus in our understanding that the limits of these two fundamental properties in humans are unknown. Making progress in gaining that insight is the aim of the work presented in this thesis.

Fine knowledge on frequency analysis and temporal coding in nonhuman species has been gained through the use of invasive techniques. Other, necessarily indirect, techniques are often applied as research tools to gain insight in hearing of humans. For peripheral hearing, these tools are mostly non-electrophysiological such as for example techniques based on psychoacoustic assessments and on evoked otoacoustic emissions (Kemp, 1978). Because of the indirect nature of these techniques, the real extents of basic properties are unknown in human and are lively debated (Ruggero and Temchin, 2005; Eustaquio-Martin and Lopez-Poveda, 2011). For example, on the one hand it is thought that frequency tuning in humans is similar to that in commonly studied laboratory mammals (Ruggero and Temchin, 2005), but on the other hand, some studies report sharper tuning (Harrison, 1981; Shera et al., 2002; Oxenham and Shera, 2003). Another neural property that is not well characterized in human and on which there is no consensus (Heinz et al., 2001; Moore, 2008; Joris and Verschooten, 2013), is neural phase-locking, the synchronization to fine-structure in sound.

In this work we studied frequency selectivity and neural phase-locking in humans with normal hearing. The focus of our research was twofold, first to determine the sharpness of neural tuning and second to obtain the frequency limit of neural phase-locking; both at the level of the auditory periphery. To achieve our goal, two paradigms were developed in which we sought to optimize stimuli, recording, and analysis. We first developed, assessed and optimized these paradigms in chinchilla and cat. Subsequently, we applied these paradigms to human volunteers with normal hearing. The desired neural contributions were extracted, analyzed and quantified. In order to make an estimate of neural properties in human at the level of single fibers, we made use of single AN fiber data of cat, chinchilla and monkey.

In the following, the results are summarized, discussed and placed in a broader context.

Frequency selectivity

We studied frequency selectivity with stimulus evoked CAPs at or near the RW using a notched-noise forward masking (NNFM) paradigm inspired by paradigms used in behavioral studies (Patterson, 1976; Oxenham and Shera, 2003). The method was first developed and assessed in cat and chinchilla and thereafter applied in macaque and human. Our ultimate goal was to quantify and compare neural frequency selectivity in human with other species.

The first important finding from an assessment of the method in chinchilla and cat was that this empirical method is a suitable means to examine frequency tuning, for a number of reasons. a) Masking of CAP responses behaved as expected from nerve physiology, with temporal recovery and masking-intensity curves similar to that of single AN fibers using tones (Harris and Dallos, 1979). b) Masking tuning curves showed a monotonic increase with notch-width, equivalent to the V-shape from tuning curve of single AN fibers. c) The sharpness of masking tuning curves, quantified as Q_{10}-factor (CAP-Q_{10}) was dependent on probe frequency and probe level. The CAP-Q_{10} increased with probe frequency with a slope similar to that of single ANFs. Moreover, at low and middle frequencies the CAP-Q_{10} followed closely the lower boundary of the single AN fibers and was only slightly higher at high frequencies. Above a certain probe level (~45 dB SPL), the CAP-Q_{10} decreased monotonously with probe level, consistent with the observations in the cochlea (Robles and Ruggero, 2001).

We applied the same CAP-method in human and macaque and found that masking behaved similar as in cat and chinchilla. We measured CAP-Q_{10} tuning in different subjects at four probe frequencies (2, 3, 4 and 6 kHz) and found again a monotonic increase in sharpness with frequency, as expected from cat and chinchilla, but with sharper tuning. Comparison between the CAP-Q_{10} trends of human and cat or chinchilla showed that CAP-tuning in human is on average ~1.6 times sharper, which is our second important finding. Note that this is not the first electrophysiological observation concerning sharper tuning in human. In a previous study of CAPs (Harrison et al., 1981), a ratio of 1.8 was reported for the relative sharpness between human an guinea pig, which has a comparable tuning as chinchilla. However, this finding seems to have received limited attention in the literature, perhaps because of a number of methodological issues (Ruggero and Temchin, 2005). The study was conducted in patients (“near normal hearing humans”); the procedure followed a clinical electrocochleographic protocol with qualitative rather than quantitative stimulus control and response assessment, the probes were filtered clicks, and the paradigm was based on simultaneous masking which is known to be confounded by 2-tone suppression.

Measurements in macaque monkey are an important component of our work, which significantly tighten the extrapolation across species. First, the measurement conditions in human were much more similar to these of monkey than to these in cat. The similarity in anatomy between monkey and human (notably the lack of a large bulla) necessitated the use of a similar trans-tympanic procedure, and similar sound levels were needed to comply with the predetermined signal-to-noise ratio criterion in the paradigm. Second, the macaque is an old-world monkey and is of course more closely related to human than to cat or chinchilla. Third, we have single fiber tuning data available for this species (Joris et al., 2011). Comparison between the CAP-Q_{10} trends of human and monkey showed that CAP-tuning in human is on average 1.35 times sharper than in monkey. This factor is not stable over the measured frequency range (2-6 kHz); it decreases with frequency from 1.6 to 1.2. This deviation comes from a clearly steeper slope of CAP-tuning of monkey relative to human and other species (Chapter 3, Fig. 6); at the highest frequencies measured, the monkey CAP-Q_{10} values approach those in humans, while at the lowest measured frequencies they are similar to those in cat and chinchilla. It is of interest to examine why the CAP-tuning of macaque has a steeper tuning-frequency slope than the other species, including human. A first and important observation is that the course of the tuning for single AN fibers in macaque (Joris et al., 2011; see also Chapter 3, Fig. 9) shows a similar pattern, converging with human Q estimates of Shera et al. (2002) and Oxenham and Shera (2003) at high frequencies but converging with cat and smaller mammals at very low frequencies. So the steeper slope in the CAP-Q_{10} trend reflects, at least partially, a real underlying trend. Nevertheless, unlike in cat and chinchilla, the CAP-Q_{10} does not parallel the neural single fiber tuning (Chapter 3, Fig. 7c vs. Fig. 7a,b). We therefore have to examine whether methodological differences between the measurements in monkey, *versus* cat and chinchilla, may underlie the steep slope measured in monkey. A possible origin of differences is cochlear compression (Chapter 1, *Cochlear amplifier and compression*), because much higher stimulus levels were needed in the monkey to compensate for the lower signal-to-noise ratio of the response. On the one hand we know from the parametric study in Chapter 2 that CAP tuning decreases with probe level (Chapter 2, Fig. 11), similar to cochlear tuning (Robles and Ruggero, 2001). On the other hand, modeling studies suggest that Q-values based on iso-response measurements can actually increase with signal level (Eustaquio-Martin and Lopez-Poveda, 2011). These two opposite effects may seem to contradict each other, but they do not because the first effect (decrease with level) is associated with compression on the probe and the second effect is associated with compression on the masker. Depending on the paradigm and its parameters (e.g. levels), both effects may occur separately or simultaneously. In the cat experiments in Chapter 2 the

stimulus levels were relatively low and masker and probe were most likely not subjected to cochlear compression, as argued by several observations (Chapter 2, Figs. 12 and 13). However, in experiments with macaque and human the levels were systematic ~30 dB higher, which means that the probe and likely also the masker (in our paradigm less stringent than the probe) were almost certainly subjected to cochlear compression. The result of a combined effect on Q_{10} is unclear, but it is conceivable that the two opposite effects counteract each other or at least to some extent. Now, if such a mutual interaction would be frequency dependent, it could affect the tuning-frequency slope of the CAP-Q_{10} trend. For example, let us assume that for low probe frequencies the probe is subjected to compression but not the masker, this would mean that for these frequencies the CAP-Q_{10} will be underestimated (as shown in Chapter 2, Fig. 11). If at high frequencies both probe and masker are subjected to compression, such an underestimation may not be present. Clearly, additional experiments are needed to investigate the effect of cochlear compression on CAP-tuning.

A related issue to the previous paragraph is why the same steep tuning-frequency slope is not present in the CAP-tuning of human, where the stimulus levels were similar to those in monkey? Perhaps even more intriguing is the shallowness of the slope of the estimates of human tuning based on behavior and emissions (Chapter 3 Fig. 9, green curves, based on Shera et al., 2002 and Oxenham and Shera, 2003), when compared to the steeper slope of the Q factors derived from CAPs and all single fiber data. Note, however, that all the primate data fit together remarkably well when taken at face value. Application of the correction function obtained from the monkey data (Chapter 3, Fig. 7c) to the human CAP data gives a prediction of human single fiber tuning (Chapter 3, blue dashed line in Fig. 8 and upper border of gray area Fig. 9) with a shallow slope and in excellent quantitative agreement with the behavioral and emission data of Shera et al. (2002) and Oxenham and Shera (2003).

Why would frequency tuning be sharper in human?

Earlier, comparative discussions of frequency tuning have looked into mechanistic (rather than teleological) factors related to sharpness of tuning (Fay, 1992). A recurring theme is that sharpness of tuning in the mammalian cochlea would depend on the ratio between the length of the basilar membrane and its frequency span (e.g. Watson, 1963; Hoppe and Langford, 1974; Harrison et al., 1981). This means that the sharper tuning in humans found in this study could be related to the longer basilar membrane than that for the other species here studied. In a limited and specific way, such an arrangement is found in some highly specialized bat species which have an “acoustic fovea” serving echolocation. The acoustic fovea is a disproportionally long area on the BM dedicated to a narrow frequency range

(Manley, 1990) which indeed exhibits sharp frequency tuning. Likewise, it seems reasonable to surmise that frequency selectivity is higher in an animal in which an octave spans a large part of basilar membrane. If this reasoning holds, the mapping of frequency on cochlear coordinates should be predictive of frequency tuning.

The mammalian cochlea has a frequency-position map that can be described by an empirical formula $f = A.(10^{ax} - k)$ (Greenwood, 1990) where f is the frequency to which a piece of membrane is tuned, A is a scaling factor which shifts the curve along the log-frequency axis; k is a constant, which introduces a low-frequency curvature to fit the "compressed" representation of low-frequencies; a is the slope (mm^{-1}) of the log-frequency-position curve at the high frequency side; and x is the independent variable, which is the distance in mm from the cochlear apex. Despite the fact that all of these parameters are species dependent, there is an inverse correlation between the slope of the log-frequency-position (a) and the length of the basilar membrane (Greenwood, 1990). For example; human, cat, chinchilla, guinea pig and monkey, have roughly the same frequency span in log units (~10 octaves), and the normalized log-frequency gradient, the product of a with the total length of the basilar membrane, is a constant and equal to 2.1. This means that for these species the log-frequency gradient (a) is inversely proportional to the length of the cochlea. So, if tuning is inversely correlated to the log-frequency gradient then we should find, in this group of species, a systematic increase in sharpness of tuning with the length of their cochlea.

To check this prediction, we compare normalized tuning/length ratios of macaque, cat, guinea pig and chinchilla with respect to human. The sharpness of neural tuning from AN fibers is obtained from the curves in Fig. 9 of chapter 3 as the average of the relative (with respect to human) tunings at 4 and 6 kHz; for human, the same measure was obtained but from the estimated neural tuning (thick black curve, Fig. 9). The relative tunings for the animals listed above are equal to 0.92 (macaque), 0.67 (cat), 0.51 (guinea pig), 0.49 (chinchilla), and 1 for human, which is here taken as the reference. The lengths of the cochleas are 25.6, 25, 18.5, 18.4 and 35 mm (taken from Liberman, 1982 and Greenwood, 1990). Normalized to human; the lengths are: 0.71, 0.71, 0.53, 0.53 and 1. From these values the normalized tuning/length ratios were calculated as 1.28 (macaque), 0.94 (cat), 0.96 (guinea pig), 0.93 (chinchilla) and 1 (human). Except for macaque, these ratios are rather consistent and close to 1. To conclude, simple reasoning based on anatomical dimensions combined with Greenwood's frequency-position function suggests that there is a relation between the sharpness of tuning and the length of the cochlea. Note that behavioral estimates of frequency analysis provide another, independent, dataset which supports this line of reasoning, as discussed by (Fay, 1992). For example, estimates of critical bandwidth span a

cochlear distance of about 1 mm in diverse species (Fay, 1992). Thus, in multiple ways, our tuning estimates derived from the CAP data are generally in line with evidence obtained completely independently. .

Neural Phase-locking

To measure neural phase-locked potentials in human, we first developed a technique to derive a measure for the neural phase-locked component (neurophonic) in the evoked cochlear potential. To achieve this, different techniques were combined to separate the neurophonic from the non-phase-locked (e.g. CAP) and non-neural phase-locked components (e.g. CM) in the evoked cochlear potential. The main techniques used to achieve this separation were signal averaging, signal filtering (in its broadest sense) and, most importantly, forward masking. The method was first assessed at the AN and RW in cat and thereafter validated by means of TTX. The result of this assessment brings us to the third important finding in this study: that there is a considerable neural phase-locked contribution present in the potential measured at the RW, which can be isolated and quantified. Moreover, we found that the neurophonic at the RW is in many ways very similar to that measured at the AN. The only noteworthy differences were a typically larger magnitude and a small time lag between the RW and AN.

The presence of a neural contribution at the RW has been reported before (Henry, 1995; He et al., 2012; Lichtenhan et al., 2013), but in our study we demonstrate that it is possible to obtain a measure for the neurophonic at the RW without altering the normal activity in the cochlear, such as neural or other blocking agents. This enabled us to determine the upper detectable frequency limit of neural phase-locking. For cat we obtained a frequency limit at the AN of 4.2 kHz and at the RW of 4.7 kHz, which is comparable to that for AN fibers (~ 5 kHz; Johnson, 1980; Joris et al., 1994) and unlike the limit reported by Snyder and Schreiner (1984) and Lichtenhan et al. (2013). Snyder and Schreiner (1984) measured phase-locking at the AN and suggested that neural phase-locking would exceed the reported limit of single AN fibers. We believe that, despite their differential recording technique, the transfer functions at the highest intensities were contaminated by stray-CM. On the other hand Lichtenhan et al. (2013) expressed their skepticism and stated that potentials measured near the cochlea would not be suitable to determine the upper frequency limit of phase-locking in humans. We showed here that it is possible to estimate the limit for phase-locking from such potentials, even when recorded at the RW, and that forward masking is key in order to remove the coherent and dominant CM at the frequency limit. It follows that the ongoing 2nd harmonic component of the neurophonic, called the “auditory nerve overlapped

waveform" by Lichtenhan et al. (2013) and mostly used as an indicator of neural activity at low frequencies, is not suitable to determine the limit.

We applied the same paradigm in human volunteers with normal hearing, but with recordings at the cochlear promontory or in the niche of the RW. These recordings were in fact interleaved with the CAP recordings obtained to estimate frequency selectivity. We also applied the same technique in macaque monkey, for which single AN fiber recordings are available in our laboratory (unpublished). We found in many respect similarities between the results from human, monkey and those obtained at the RW in cat. In all species, the neurophonic showed similar temporal and spectral properties, such as adaptation, masking, a delay with respect to CM and similar spectral band-pass shapes around 0.8-1.0 kHz. The most notable difference in human or monkey compared to cat was a smaller magnitude of the neurophonic. The neurophonic was on average >30 dB smaller with an additional reduction in signal-to-noise(floor) ratio of ~20 dB (Chapter 5, Fig. 8A).

The upper detectable frequency limits of neural phase-locked contributions that we obtained were 3.3 kHz in human and 4.0 kHz in macaque. Thus, the limit in human was 0.28 octaves or 1.2 times lower than in monkey and 0.51 octaves or 1.42 times lower than in cat (4.7 kHz). The ratio between human and cat (0.51 octave) suggests that the difference is not as large as previously suggested (about 1 octave, Joris and Verschooten, 2013). The estimation in that study was based on a comparison between frequency limits of psychophysical binaural sensitivity studies in human (1.4 kHz, based on Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Schiano et al., 1986; Dunai and Hartmann, 2011) and in cat (2.8 kHz, based on Wakeford and Robinson, 1974; Jackson et al., 1996). Remark that the behavioral limit in cat is similar to that reported for physiological interaural sensitivity in the inferior colliculus (Yin and Kuwada, 1983), the thalamus (Ivarsson et al., 1988), and primary auditory cortex (Reale and Brugge, 1990), and that these limits are much lower than that of the auditory nerve (e.g. cat: 2.8 kHz versus ~5 kHz and 4.7 kHz for single AN fibers and for the neurophonic in this study, respectively). This indicates that the upper frequency of useful phase-locking in the binaural auditory system is lower than the phase-locking limit at the AN. Moreover, there are reasons to believe (Joris and Verschooten, 2013) that this would not be different for the monaural system.

The upper frequency limit in human

In cat we found good agreement between the detectable frequency limit of phase-locking at the RW (4.7 kHz) and that from the VS of single nerve-fibers (~5 kHz). A severe limitation in the data for human and monkey is the signal-to-noise(floor) ratio of neurophonic, which in human and macaque was ~20 dB lower than in cat. This in itself is a reason to believe that the limits from human and monkey are underestimated. However, the limit of 4 kHz obtained in macaque seems to reflect the roll-off of VS in that species rather well: although the upper limit of significant phase-locking in monkey is similar to cat, the slope of that roll-off differs so that phase-locking at frequencies between 2-4 kHz is actually somewhat lower than in cat (unpublished). The minimum statement we can make is that our measurements in human and macaque provide no support for the hypothesis that neural phase-locking in humans extends higher than in cat. We believe our data argue that the limit in humans is actually lower, though not fully by an octave as we proposed earlier (Joris and Verschooten, 2013).

General Conclusion

In this thesis two methods have been developed based on forward masking and stimulus evoked neural mass potential to assess two fundamental neural properties at the level of the periphery in human. Based on the results in different species, we can minimal state that there is no support for the hypothesis that humans have a higher upper frequency limit of phase-locking than cats, and secondly that our results support the hypothesis that frequency tuning in humans is sharper than in commonly studied animals, such as chinchilla and cat.

Perspectives and future research

In experiments with macaque and human the probe and masker levels were systematically higher than in cat. From Chapter 2 we know that sharpness of CAP-tuning is negatively correlated with probe level, but little is known about the effect of masker level or the combined effect on CAP-tuning. It has recently been suggested (Eustaquio-Martin and Lopez-Poveda, 2011) that Q-factors based on iso-response measurements can be overestimated when using large signal levels. In this thesis we found no direct evidence of such systematic overestimation, also not in monkey (Chapter 3, Fig.7) where levels were larger than in cat, but similar to those in human experiments. However, from the point of view of cochlear compression, it is not unreasonable to think that the effect of compression on the level (e.g. masker) has a tendency to sharpen up iso-response tuning slopes and therefore also

Q_{10}. On the other hand, preliminary modeling simulating the effect of the broadening of the masker's excitation pattern – concomitant with the compression of level – indicates an opposite tendency, i.e. broadening the tuning. In a first electrophysiological experiment in cat we started to investigate the CAP-tuning dependency on masker level, but these results were not stable enough to draw a definitive conclusion. Therefore, more data is needed in order to answer this debated topic.

In Chapter 4 we showed the effect of probe level on the magnitude and the relative phase-delay of the CM and decaying neurophonic. This kind of data is in many aspects interesting for further analysis, especially when it is extended to higher levels and to a larger range of probe frequencies. First, the dependency of the upper frequency limit of phase-locking on probe level can be examined and the maximum frequency limit of phase-locking can be determined, similar to the maximum vector strength of single AN fibers. Second, it would allow us to verify if the maximum frequency limit would exceed this of single AN fibers. This was previously suggested by Snyder and Schreiner (1984, their Figure 12; limit >10 kHz for probe levels of 70 and 80 dB SPL), but we had strong arguments (Chapter 4) that these transfer functions are contaminated by stray-CM. Third, additional knowledge of the phase delay, group delay and dispersion of the different phase-locked contributions at different probe levels can give us better understanding about the origin and relative timing of (neural) generators involved in the potentials. Moreover, these results would help us to extend and/or improve existing computational peripheral auditory models (e.g., from the Carney Lab). This research is in progress.

The experimental procedure in humans developed in this study gives the opportunity to study other auditory systems involved in the peripheral auditory pathway of normal-hearing humans. An interesting system that is not completely understood is the olivocochlear efferent system. This system allows some control over the cochlea and has the ability to suppress or enhance AN fiber responses (Galambos, 1956). The olivocochlear system consists of two subsystems (Warr and Guinan, 1979), the medial olivocochlear system, which terminates onto the cochlear OHCs, and the lateral olivocochlear system, which innervates afferent AN fibers from IHCs. The medial olivocochlear system is rather well studied and thought to be involved in the facilitation of the detection or discrimination of sounds in the presence of background noise, whereas little is known about the function of the lateral olivocochlear system. The combination of the experimental procedure in human with a modified (forward masked) paradigm tailored to the olivocochlear system would allow us to extend our knowledge of the olivocochlear system in human.

The procedure and paradigm to temporally separate neurophonic and microphonic are not only of scientific interest, but could also have a clinical significance. For instance, it could be used in the diagnosis of peripheral auditory neural diseases, such as auditory neuropathy where OHC function is normal but neural transmission to the auditory nerve and brainstem is impaired. It might also be of interest for the diagnosis of pulsatile tinnitus in case of vascular compression of the auditory nerve. To conclude, further research into this topic could lead to the development of a clinical tool for the diagnosis of several peripheral auditory disorders.

REFERENCES

Allen JB (1983) Magnitude and phase-frequency response to single tones in the auditory nerve. The Journal of the Acoustical Society of America 73:2071-2092.

Antoli-Candela EJ, Kiang NYS (1978) Unit activity underlying the N1 potential. In: Evoked Electrical Activity in the Auditory Nervous System (Naunton R, Fernandez C, eds), pp 165-189. New York: Academic Press.

Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. Nature 451:197-201.

Blauert J (1983) Spatial Hearing. Cambridge: MIT Press.

Brownell WE, Bader CR, Bertrand D, de Ribaupierre Y (1985) Evoked mechanical responses of isolated cochlear outer hair cells. Science 227:194-196.

Carlier E, Lenoir M, Pujol R (1979) Development of cochlear frequency selectivity tested by compound action potential tuning curves. hear Res 1:197-201.

Carney LHC (1994) Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness. HearRes 76:31-44.

Carney LHC, Yin TCT (1988) Temporal Coding of Resonances By Low-Frequency Auditory Nerve Fibers: Single-Fiber Responses and a Population Model. J Neurophysiol 60:1653-1677.

Chadwick RS (1998) Compression, gain, and nonlinear distortion in an active cochlear model with subpartitions. Proceedings of the National Academy of Sciences 95:14594-14599.

Chimento TC, Schreiner CE (1990) Time course of adaptation and recovery from adaptation in the cat auditory-nerve neurophonic. J Acoust Soc Am 88:857-864.

Crawford AC, Fettiplace R (1985) The mechanical properties of ciliary bundles of turtle cochlear hair cells. The Journal of physiology 364:359-379.

Dallos P (1973) The auditory periphery. New York: Academic Press.

Dallos P, Cheatham MA (1976) Compound action potential (AP) tuning curves. J Acoust Soc Am 59:591-597.

Dallos P, Evans BN (1995) High-frequency motility of outer hair cells and the cochlear amplifier. Science 267:2006-2009.

Dallos P, Wu X, Cheatham MA, Gao J, Zheng J, Anderson CT, Jia S, Wang X, Cheng WH, Sengupta S, He DZ, Zuo J (2008) Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. Neuron 58:333-339.

Davis H (1983) An active process in cochlear mechanics. Hear Res 9:79-90.

Dolan DF, Nuttall AL, Avinash G (1990) Asynchronous neural activity recorded from the round window. The Journal of the Acoustical Society of America 87:2621-2627.

Dolan DF, Xi L, Nuttall AL (1989) Characterization of an EPSP-like potential recorded remotely from the round window. J Acoust Soc Am 86:2167-2171.

Dunai L, Hartmann WM (2011) Frequency dependence of the interaural time difference thresholds in human listeners. J Acoust Soc Am 129:2485–2485.

Eggermont JJ (1976) Electrophysiological study of the normal and pathological human cochlea. II. Neural responses. Rev Laryngol Otol Rhinol (Bord) 97 Suppl:497-506.

Eggermont JJ (1976b) Electrocochleography. In: Handbook of sensory physiology (Keidel WD, Neff WD, eds), pp 625-705. Berlin: Springer.

Eggermont JJ (1977) Compound actionpotential tuning curves in normal and pathological human ears. J Acoust Soc Am 62:1247-1251.

Eustaquio-Martin A, Lopez-Poveda EA (2011) Isoresponse versus isoinput estimates of cochlear filter tuning. J Assoc Res Otolaryngol 12:281-299.

Evans EF (1972) The frequency response and other properties of single fibres in the guinea-pig cochlear nerve. J Physiol 226:263-287.

Faingold CL, Boersma A, Caspary DM (1991) Involvement of GABA in acoustically-evoked inhibition in inferior colliculus neurons. Hear Res 52:201-216.

Fay R (1992) Structure and Function in Sound Discrimination Among Vertebrates. In: The Evolutionary Biology of Hearing (Webster D, Popper A, Fay R, eds), pp 229-263: Springer New York.

Galambos R (1956) Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea. Journal of neurophysiology 19:424-437.

Gibson WPR, Moffat DA, Ramsden RT (1977) Clinical electrocochleography in the diagnosis and management of Ménière's disorder. Audiology 16:389-401.

Glasberg BR, Moore BC (1990) Derivation of auditory filter shapes from notched-noise data. Hear Res 47:103-138.

Glasberg BR, Moore BC, Patterson RD, Nimmo-Smith I (1984) Dynamic range and asymmetry of the auditory filter. J Acoust Soc Am 76:419-427.

Goldberg JM, Brown PB (1969) Response of Binaural Neurons of Dog Superior Olivary Complex to Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. J Neurophysiol 22:613-636.

Goldstein MH, Kiang NYS (1958) Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. The Journal of the Acoustical Society of America 30:107-114.

Greenwood DD (1990) A cochlear frequency-position function for several species--29 years later. J Acoust Soc Am 87:2592-2605.

Harris DM (1979) Action potential suppresion, tuning curves and thresholds: comparison with single fiber data. Hear Res 1:133-154.

Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. Journal of neurophysiology 42:1083-1107.

Harrison RV (1981) Rate-versus-intensity functions and related AP responses in normal and pathological guinea pig and human cochleas. The Journal of the Acoustical Society of America 70:1036-1044.

Harrison RV, Aran JM, Erre JP (1981) AP tuning curves from normal and pathological human and guinea pig cochleas. J Acoust Soc Am 69:1374-1385.

He W, Porsov E, Kemp D, Nuttall AL, Ren T (2012) The group delay and suppression pattern of the cochlear microphonic potential recorded at the round window. PLoS One 7:e34356.

Heinz MG, Colburn HS, Carney LHC (2001) Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. Neural Comput 13:2273-2316.

Henry KR (1995) Auditory nerve neurophonic recorded from the round window of the Mongolian gerbil. Hear Res 90:176-184.

Hoppe SA, Langford TL (1974) Binaural interaction in cat and man. I. Signal detection and noise cross correlation. J Acoust Soc Am 55:1263-1265.

Ivarsson C, De Ribaupierre Y, De Ribaupierre F (1988) Influence of auditory localization cues on neuronal activity in the auditory thalamus of the cat. Journal of neurophysiology 59:586-606.

Jackson LL, Heffner HE, Heffner RS (1996) Species differences in the upper limit of binaural phase discrimination. Assoc Res Otolaryngol Abstr 19:63.

Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J Acoust Soc Am 68:1115-1122.

Joris PX (2010) Het snelle slakkenhuis. In: Lessen voor de eenentwintigste eeuw - In gesprek met morgen (Raymaekers B, ed), pp 173-192. Leuven: Universitaire Pers Leuven.

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proceedings of the National Academy of Sciences of the United States of America 108:17516-17520.

Joris PX, Carney LH, Smith PH, Yin TC (1994b) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. Journal of neurophysiology 71:1022-1036.

Joris PX, Smith PH, Yin TC (1994a) Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. Journal of neurophysiology 71:1037-1051.

Joris PX, Verschooten E (2013) On the limit of neural phase-locking to fine-structure in humans. Basic Aspects of Hearing: Physiology and Perception Editors: BCJ Moore, RD Patterson, I Winter, RP Carlyon, and HE Gockel (Springer, NY), in press.

Kemp DT (1978) Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 64:1386-1391.

Kiang NYS (1984) Peripheral neural processing of auditory information. In: Handbook of Physiology. Section 1, Volume III, part Edition (Brookhart JM, Mountcastle VB, eds), pp 639-674. Bethesda: American Physiological Society.

Kiang NYS, Moxon EC, Kahn AR (1976) The Relationship of Gross Potentials Recorded fron the Cochlea to Single Unit Activity in the Auditory Nerve. In: Electrocochleography (Ruben RJ, Elberling C, Salomon G, eds), pp 95-115. Baltimore: University Park Press.

Kiang NYS, Watanabe T, Thomas EC, Clark LF (1965) Discharge patterns of single fibers in the cat's auditory nerve, 35 Edition. MIT Press, Cambridge: Research Monograph No 35.

Kidd RC, Weiss TF (1990) Mechanisms that degrade timing information in the cochlea. Hear Res 49:181-208.

Kim DO, Molnar CE (1979) A Population Study of Cochlear Nerve Fibers: Comparison of Spatial Distributions of Average-Rate and Phase-Locking Measures of Responses to Single Tones. J Neurophysiol 42:16-30.

Klumpp R, Eady H (1956) Some measurements of interaural time differences thresholds. J Acoust Soc Am 28:859-864.

Kraus N, Bradlow AR, Cheatham MA, Cunningham J, King CD, Koch DB, Nicol TG, McGee TJ, Stein LK, Wright BA (2000) Consequences of neural asynchrony: a case of auditory neuropathy. Journal of the Association for Research in Otolaryngology : JARO 1:33-45.

Levine RA, Gardner JC, Fullerton BC, Stufflebeam SM, Carlisle EW, Furst M, Rosen BR, Kiang NY (1993) Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials. Hear Res 68:73-88.

Liberman MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am 63:442-455.

Liberman MC (1982) The cochlear frequency map for the cat: labeling auditory-nerve fibers of known characteristic frequency. J Acoust Soc Am 72:1441-1449.

Liberman MC, Gao J, He DZ, Wu X, Jia S, Zuo J (2002) Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. Nature 419:300-304.

Lichtenhan JT, Cooper NP, Guinan JJ, Jr. (2013) A new auditory threshold estimation technique for low frequencies: proof of concept. Ear Hear 34:42-51.

Lopez-Poveda EA, Barrios LF, Alves-Pinto A (2007) Psychophysical estimates of level-dependent best-frequency shifts in the apical region of the human basilar membrane. J Acoust Soc Am 121:3646-3654.

Manley GA (1990) Peripheral hearing mechanisms in reptiles and birds: Springer-Verlag.

McGee T, Ryan A, Dallos P (1976) Psychophysical tuning curves of chinchillas. J Acoust Soc Am 60:1146-1150.

McMahon CM, Patuzzi RB (2002) The origin of the 900 Hz spectral peak in spontaneous and sound-evoked round-window electrical activity. Hearing research 173:134-152.

Michelet P, Mc Laughlin M, van der Heijden M, Joris P (2011) Synchronization to pure and amplitude-modulated tones in the auditory nerve of macaque monkey. Assoc Res Otolaryngol Abstract 34:223.

Moller AR (1984) Pathophysiology of tinnitus. The Annals of otology, rhinology, and laryngology 93:39-44.

Moller AR (1986) Effect of click spectrum and polarity on round window N1N2 response in the rat. Audiology 25:29-43.

Moller AR, Jho HD (1989) Response from the exposed intracranial human auditory nerve to low-frequency tones: basic characteristics. Hear Res 38:163-175.

Moore BC (2008) The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. J Assoc Res Otolaryngol 9:399-406.

Nelson DA, Schroder AC, Wojtczak M (2001) A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 110:2045-2064.

Niemiec AJ, Yost WA, Shofner WP (1992) Behavioral measures of frequency selectivity in the chinchilla. J Acoust Soc Am 92:2636-2649.

Oxenham AJ, Plack CJ (1997) A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. J Acoust Soc Am 101:3666-3675.

Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J Assoc Res Otolaryngol 4:541-554.

Oxenham AJ, Simonson AM (2006) Level dependence of auditory filters in nonsimultaneous masking as a function of frequency. J Acoust Soc Am 119:444-453.

Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair cells. Hear Res 24:1-15.

Palmer AR, Shackleton TM (2009) Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency. Journal of the Association for Research in Otolaryngology : JARO 10:233-250.

Patterson RD (1976) Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 59:640-654.

Patterson RD, Nimmo-Smith I (1980) Off-frequency listening and auditory-filter asymmetry. The Journal of the Acoustical Society of America 67:229-245.

Patterson RD, Nimmo-Smith I, Weber DL, Milroy R (1982) The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold. J Acoust Soc Am 72:1788-1803.

Patuzzi RB, Brown DJ, McMahon CM, Halliday AF (2004) Determinants of the spectrum of the neural electrical activity at the round window: transmitter release and neural depolarisation. Hear Res 190:87-108.

Peake WT, Kiang NY (1962) Cochlear responses to condensation and rarefaction clicks. Biophysical journal 2:23-34.

Peake WT, Rosowski JJ (1991) Impedance matching, optimum velocity, and ideal middle ears. Hear Res 53:1-6.

Pickles JO (1979) Psychophysical frequency resolution in the cat as determined by simultaneous masking and its relation to auditory-nerve resolution. J Acoust Soc Am 66:1725-1732.

Plack CJ, Oxenham AJ (1998) Basilar-membrane nonlinearity and the growth of forward masking. J Acoust Soc Am 103:1598-1608.

Portmann M, Aran JM (1971) Electrocochleography. Laryngoscope 81:899-910.

Prijs VF (1980) In: The single-fibre representation in the compound action potential of the auditory nerve, p 91: PhD thesis.

Prijs VF (1986) Single-unit response at the round window of the guinea pig. Hearing research 21:127-133.

Rasmussen AT (1943) Outlines of neuro-anatomy. Dubuque (Iowa): Wm. C. Brown.

Reale RA, Brugge JF (1990) Auditory cortical neurons are sensitive to static and continuously changing interaural phase cues. J Neurophysiol 64:1247-1260.

Recio-Spinoso A, Temchin AN, van Dijk P, Fan YH, Ruggero MA (2005) Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. Journal of neurophysiology 93:3615-3634.

Relkin EM, Smith RL (1991) Forward masking of the compound action potential: thresholds for the detection of the N1 peak. Hear Res 53:131-140.

Relkin EM, Turner CW (1988) A reexamination of forward masking in the auditory nerve. J Acoust Soc Am 84:584-591.

Rhode WS, Roth GL, Recio-Spinoso A (2010) Response properties of cochlear nucleus neurons in monkeys. Hear Res 259:1-15.

Rhode WS, Smith PH (1985) Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers. Hear Res 18:159-168.

Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. Physiological reviews 81:1305-1352.

Rose JE, Brugge JF, Anderson DJ, Hind JE (1967) Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J Neurophysiol 30:769-793.

Rose JE, Brugge JF, Anderson DJ, Hind JE (1968) Patterns of activity in single auditory nerve fibres of the squirrel monkey. In: de Reuck, AVS, Knight, J (Eds), Ciba Foundation Symposium on Hearing Mechanisms in Vertrebrates J&A Churchill, London, pp144–157.

Rose JE, Hind JE, Anderson DJ, Brugge JF (1971) Some Effects of Stimulus Intensity on Response of Auditory Nerve Fibers in the Squirrel Monkey. Journal of Neurophysiology 34:685-699.

Ruggero MA, Rich NC (1987) Timing of Spikes initiation in Cochlear Afferents: Dependence on Site of Innervation. J Neurophysiol 58:379-403.

Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L (1997) Basilar- membrane responses to tones at the base of the chincilla cochlea. J Acoust Soc Am 101:2151-2163.

Ruggero MA, Temchin AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. Proceedings of the National Academy of Sciences of the United States of America 102:18614-18619.

Russel AD (2004) Introduction to Psychology: Wadsworth Publishing.

Russell I, Palmer A (1986) Filtering due to the inner hair-cell membrane properties and its relation to the phase-locking limit in cochlear nerve fibres. In: Auditory Frequency Selectivity (Moore BCJ, Patterson RD, eds), pp 199-207. New York: Plenum Press.

Salvi RJ, Giraudi DM, Henderson D, Hamernik RP (1982) Detection of sinusoidally amplitude modulated noise by the chinchilla. J Acoust Soc Am 71:424-429.

Salvi RJ, Henderson D, Hamernik RP (1979) Single auditory nerve fiber and action potential latencies in normal and noise treated chinchillas. Hear Res 1:237-251.

Sayles M, Winter IM (2010) Equivalent-rectangular bandwidth of single units in the anaesthetized guinea-pig ventral cochlear nucleus. Hear Res 262:26-33.

Schiano JL, Trahiotis C, Bernstein LR (1986) Lateralization of low-frequency tones and narrow bands of noise. J Acoust Soc Am 79:1563-1570.

Searchfield GD, Munoz DJ, Thorne PR (2004) Ensemble spontaneous activity in the guinea-pig cochlear nerve. Hear Res 192:23-35.

Shera CA, Guinan JJ, Jr. (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am 105:782-798.

Shera CA, Guinan JJ, Jr. (2003) Stimulus-frequency-emission group delay: a test of coherent reflection filtering and a window on cochlear tuning. Journal of the Acoustical Society of America 113:2762-2772.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proceedings of the National Academy of Sciences of the United States of America 99:3318-3323.

Shera CA, Guinan JJ, Jr., Oxenham AJ (2010) Otoacoustic estimation of cochlear tuning: validation in the chinchilla. J Assoc Res Otolaryngol 11:343-365.

Snyder RL, Schreiner CE (1984) The auditory neurophonic: basic properties. Hearing research 15:261-280.

Snyder RL, Schreiner CE (1985) Forward masking of the auditory nerve neurophonic (ANN) and the frequency following response (FFR). Hear Res 20:45-62.

Spagnoli SD, Saunders JC (1987) Threshold sensitivity and frequency selectivity measured with round window whole nerve action potentials in the awake, restrained chinchilla. Otolaryngol Head Neck Surg 96:99-105.

Starr A, Sininger Y, Nguyen T, Michalewski HJ, Oba S, Abdala C (2001) Cochlear receptor (microphonic and summating potentials, otoacoustic emissions) and auditory pathway (auditory brain stem potentials) activity in auditory neuropathy. Ear Hear 22:91-99.

Tasaki I (1954) Nerve impulses in individual auditory nerve fibers of guinea pig. J Neurophysiol 17:97-122.

Teich M, Khanna S, Guiney P (1993) Spectral characteristics and synchrony in primary auditory-nerve fibers in response to pure-tone acoustic stimuli. J Stat Phys 70:257-279.

Temchin AN, Rich NC, Ruggero MA (2008) Threshold tuning curves of chinchilla auditory-nerve fibers. I. Dependence on characteristic frequency and relation to the magnitudes of cochlear vibrations. J Neurophysiol 100:2889-2898.

Temchin AN, Ruggero MA (2010) Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics. J Assoc Res Otolaryngol 11:297-318.

Van Heusden E, Smoorenburg GF (1981) Eighth-nerve action potentials evoked by tone bursts in cats before and after inducement of an acute noise trauma. Hear Res 5:1-23.

Verschooten E, Robles L, Kovacic D, Joris PX (2012) Auditory nerve frequency tuning measured with forward-masked compound action potentials. Journal of the Association for Research in Otolaryngology : JARO 13:799-817.

Versnel H, Prijs VF, Schoonhoven R (1992b) Round-window recorded potential of single-fibre discharge (unit response) in normal and noise-damaged cochleas. Hearing research 59:157-170.

Versnel H, Schoonhoven R, Prijs VF (1992a) Single-fibre and whole-nerve responses to clicks as a function of sound intensity in the guinea pig. Hearing research 59:138-156.

von Békésy G (1960) Experiments in hearing. New York: Mc Graw Hill.

von Békésy G, Rosenblith WA (1951) In S. S. Stevens (Ed.). The mechanical properties of the ear. New York: Wiley.

Voss SE, Rosowski JJ, Peake WT (1996) Is the pressure difference between the oval and round windows the effective acoustic stimulus for the cochlea? J Acoust Soc Am 100:1602-1616.

Wakeford OS, Robinson DE (1974) Lateralization of tonal stimuli by the cat. The Journal of the Acoustical Society of America 55:649-652.

Wang B (1979) The relation between the compound action potential and unit discharges of the auditory nerve. Doctoral dissertation, Dept of Electrical Engineering and Computer Sciences, MIT.

Warr WB, Guinan JJ, Jr. (1979) Efferent innervation of the organ of corti: two separate systems. Brain Res 173:152-155.

Watson CS (1963) Masking of tones by noise for the cat. The Journal of the Acoustical Society of America 35:167-173.

Weinberger MM, Kitzes LM, D.A. G (1970) Some characteristics of the auditory neurophonic. Experientia (Basle) 26:46-48.

Weiss TF, Rose C (1988a) A comparison of synchronization filters in different auditory receptor organs. Hear Res 33:175-180.

Weiss TF, Rose C (1988b) Stages of degradation of timing information in the cochlea: A comparison of hair-cell and nerve-fiber responses in the alligator lizard. Hear Res 33:167-174.

Westerman LA, Smith RL (1984) Rapid and short-term adaptation in auditory nerve responses. Hear Res 15:249-260.

Wever EG, Bray CW (1930) Action Currents in the Auditory Nerve in Response to Acoustical Stimulation. Proc Natl Acad Sci U S A 16:344-350.

Yates GK, Johnstone BM, Patuzzi RB, Robertson D (1992) Mechanical preprocessing in the mammalian cochlea. Trends in neurosciences 15:57-61.

Yin TCT, Kuwada S (1983) Binaural Interaction in Low-Frequency Neurons in Inferior Colliculus of the Cat. II. Effects of Changing Rate and Direction of Interaural Phase. J Neurophysiol 50:1000-1018.

Zwislocki JJ, Feldman RS (1956) Just noticeable differences in dichotic phase. J Acoust Soc Am 28:860-864.

Summary

The ability to hear is very important for humans and probably all mammals. It is important for predators to detect their prey, for prey to detect their predators, and in many animals it allows advanced communication, which is particularly important in humans. The sound that reaches our ears is typically a quite complex mixture of different sources. One of the capabilities of our auditory system is that it can analyze and segregate this mixture into its underlying sources. For example, one is able to follow a particular voice at a cocktail party even though other voices and background sounds are present. Two primary processes involved are the ability to resolve different spectral components in complex sounds and the ability to encode the temporal fine-structure of the waveform in the sound. These two fundamental properties are reflected in the neural signal of auditory nerve fibers (AN fibers) as a place-rate code and time code and are referred to as frequency selectivity and neural phase-locking. Frequency selectivity is often degraded in clinical conditions, which affects hearing in virtually all its capabilities, but the exact quantitative nature of the normal "filters" is unknown in human and is lively debated. Neural phase-locking is necessary for low-frequency sound localization in azimuth (left-right), but its role in other aspects of perception has been discussed for more than a century and remains to this day a contentious issue. One of the key questions in this debate is whether or not useful phase-locking in human extends to much higher frequencies (8-10 kHz) than that found in common laboratory animals such as cat (~5 kHz). The aim of this thesis was to determine the limits of these two fundamental properties, frequency selectivity and phase-locking, in the human auditory nerve, which is the bottleneck that feeds the central auditory nervous system.

Frequency analysis and temporal coding in the auditory nerve of nonhuman species can be studied with electrophysiological techniques in which individual neurons are recorded. Unfortunately, these techniques are invasive and are not applicable in humans. To gain insight in the human peripheral hearing, other indirect techniques are available, which are mostly non-electrophysiological, but there is no consensus regarding their interpretation in terms of cochlear and/or auditory nerve physiology. In order to assess frequency selectivity and neural phase-locking at the level of the auditory nerve in humans, we developed new stimuli, recording, and analysis techniques to electrophysiologically quantify these properties from cochlear mass potentials in humans.

The first two studies in this thesis examined frequency selectivity in chinchilla, cat, monkey, and human based on stimulus evoked compound action potentials (CAPs) using a notched-noise forward masking paradigm. The method was first developed and assessed in cat and chinchilla and thereafter applied in macaque and human. These two studies have led to the following findings.

A) The empirical method that was developed and assessed in cat was a suitable means to examine frequency tuning for a number of reasons. a) Masking of CAP responses behaved as expected from the physiology of single nerve fibers, with temporal recovery and masking-intensity curves similar to these fibers when using tones. b) Masking tuning curves showed a monotonic increase with notch-width, equivalent to the V-shape from tuning curve of single auditory nerve fibers. c) The sharpness of masking tuning curves, quantified as a Q_{10}-factor (CAP-Q_{10}) was dependent on probe frequency and probe level. The CAP-Q_{10} increased with probe frequency with a slope similar to that of single auditory nerve fibers. Moreover, at low and middle frequencies the CAP-Q_{10} followed closely the lower boundary of single nerve fibers and was only slightly higher at high frequencies. Above a certain probe level (~45 dB SPL), the CAP-Q_{10} decreased monotonically with probe level, consistent with cochlea physiology.

B) Measurements with the same CAP-method in human and macaque behaved similar to cat and chinchilla, but the responses and signal-to-noise(floor) ratios were much smaller and therefore higher stimulus levels (>30 dB) were needed. CAP tuning in human was 1.6 times sharper than in cat and chinchilla and on average 1.35 times sharper than in macaque monkey.

C) Estimated neural tuning in human, obtained with the average correction function based on the other species in this study, shows that it is sharper than that measured from single auditory nerve fibers in chinchilla, cat, and macaque. Furthermore, based on only the correction function of macaque, human neural tuning is in excellent quantitative agreement with behavioral and emission data.

The last two studies in this thesis examined neural phase-locking in cat, monkey, and human based on ongoing stimulus evoked potentials. To achieve this, different techniques were combined to separate the neurophonic from the non-phase-locked (e.g., CAP) and non-neural phase-locked components (e.g., cochlear microphonic) in the evoked potential; the most important technique applied was forward masking. The method was first assessed at the auditory nerve and round window in cat and thereafter validated by means of Tetrodotoxin. These two studies led to the following findings.

A) The potential near the round window has a considerable neural phase-locked contribution that can be isolated and quantified with the present method. We found that this component is in many ways similar to the component measured on the auditory nerve itself. The only noteworthy differences were a typically larger magnitude and a smaller time lag at the round window than at the auditory nerve. The frequency limit we obtained with this method in cat was 4.7 kHz, which is very similar to the limit reported for individual nerve fibers (~5 kHz).

B) There were many similarities between the measurements in human, monkey and those at the round window in cat, with similar temporal and spectral properties. The most notable difference in human and monkey compared to cat was a much smaller magnitude (-30 dB) and also signal-to-noise(floor) ratio for the neurophonic.

C) The upper detectable frequency limits of neural phase-locking obtained from stimulus evoked potentials was 3.3 kHz in human and 4.0 kHz in macaque. Data of single AN fibers in macaque show a higher upper limit that is ~1.2 times higher than 4 kHz. This suggests that the limit of single AN fibers in human is also slightly higher than obtained in this study; somewhere between 3.3 and 4.0 kHz.

To summarize, the minimum statement we can make based on the results in this thesis is, first, that there is no support for the hypothesis that humans have a higher upper frequency limit of phase-locking than cat and, second, that our results support the hypothesis that humans have sharper frequency tuning than commonly studied animals.

Samenvatting

Het vermogen om te horen is zeer belangrijk voor mens en dier. Het gehoor laat ons toe om met elkaar te communiceren en is tevens voor veel diersoorten van levensbelang. Het geluid dat onze oren bereikt, is meestal een complex mengsel afkomstig van verschillende geluidsbronnen. Een van de mogelijkheden die ons auditieve systeem heeft is dat het in staat is om de akoestische stroom van informatie te analyseren en om de onderliggende geluidsbronnen van elkaar te onderscheiden. Bijvoorbeeld, op een cocktailparty zijn mensen in staat om de spraak van een bepaalde persoon te volgen terwijl er nog andere stemmen en doordringende achtergrondgeluiden aanwezig zijn. Twee primaire processen die hiervoor belangrijk zijn, zijn het vermogen om spectrale componenten in complexe geluiden van elkaar te scheiden en het vermogen om de fijne tijdstructuur van de golfvorm van het geluid door te geven. Deze twee fundamentele eigenschappen zijn terug te vinden in de neurale signalen van auditieve zenuwvezels onder de vorm van de zgn. plaats- en tijdscode; zij worden ook frequentie selectiviteit en neurale fasekoppeling genoemd. Frequentie selectiviteit is vaak gedegradeerd in klinische condities, en dit tast vrijwel alle vermogens van het gehoor aan, maar de exacte kwantitatieve aard van de normale “filters” zijn niet gekend bij de mens en dit is een heet hangijzer. Neurale fasekoppeling is noodzakelijk voor laagfrequente azimutale geluidslokalisatie (links-rechts); de rol in andere aspecten van de geluidsperceptie wordt al meer dan een eeuw bediscussieerd maar blijft tot op heden een omstreden onderwerp. Een van de belangrijkste vragen bij dit debat is of de bruikbare fasekoppeling bij de mens al dan niet veel hoger is (bijv. 8-10 kHz) dan bij veelgebruikte laboratoriumdieren zoals de kat (~5 kHz). Dit proefschrift heeft als doel om de grenzen van deze twee fundamentele eigenschappen bij de mens te bepalen.

Frequentieanalyse en temporele codering in de gehoorzenuw van niet-menselijke diersoorten kan bestudeerd worden met elektrofysiologische technieken. Helaas zijn deze technieken zeer invasief en niet direct toepasbaar bij de mens. Om inzicht te krijgen in de werking van het gehoor bij de mens zijn er een aantal indirecte technieken, die voornamelijk niet elektrofysiologisch zijn, maar er is geen consensus aangaande hun interpretatie in termen van de fysiologie van het slakkenhuis en de zenuw. Teneinde frequentieselectiviteit en neurale fasekoppeling op het perifere niveau van de gehoorzenuw te beoordelen, ontwikkelden we nieuwe stimuli, opnametechnieken en analysetechnieken om deze eigenschappen elektrofysiologisch te kwantificeren uit massa potentialen afkomstig van het slakkenhuis.

In de eerste twee studies van dit proefschrift werd de frequentieselectiviteit in de chinchilla, kat, aap en mens bestudeerd met behulp van een “notched-noise forward masking”

paradigma en stimulus geëvoceerde massa potentialen (CAPs). Deze methode werd eerst ontwikkeld en geëvalueerd in katten en chinchilla's en daarna toegepast in de aap (makaak) en in goed horende vrijwilligers. Deze studies hebben geleid tot volgende bevindingen:

A) Onderzoek met de empirische methode ontwikkeld in de kat toonde aan dat de methode geschikt is om de frequentieselectiviteit van individuele auditieve zenuwvezels in te schatten, en dit om volgende redenen. a) Het gedrag van CAP reacties op voorwaartse maskering was in goede overeenstemming met fysiologische responsen van individuele zenuwvezels op tonen: het temporeel herstel t.g.v. maskering was in overeenstemming met de gemeten maskering-intensiteits curven. b) De afstemmings-curven bekomen met de methode van voorwaartse maskering vertoonden een monotoon stijgend verloop met notch-breedte, equivalent met de V-vorm van drempel-afstemmings-curven bekomen van individuele zenuwvezels. c) De scherpte van de met voorwaartse maskering bekomen afstemmings-curven, hier gekwantificeerd als een Q_{10}-factor (CAP-Q_{10}), was zowel afhankelijk van de sonde frequentie als het sonde niveau. De scherpte van de CAP-Q_{10} waarden verhoogt met sonde frequentie, waarbij de bijhorende helling vergelijkbaar was met die van individuele zenuwvezels. Bij lage en middenfrequenties was de waarde van de CAP-Q_{10} kleiner maar gelijk aan de ondergrens van individuele zenuwvezels; bij de hogere frequenties was deze slechts iets hoger. De CAP-Q_{10} waarde verminderde, boven een bepaalde drempel (~45 dB SPL), monotoon met het sonde niveau, en dit in overeenstemming met de fysiologie van het slakkenhuis.

B) De CAP-experimenten in mens en aap (makaak) waren in goede overeenstemming met deze van kat en chinchilla. Echter de massa potentialen en hun signaal-tot-ruis(vloer) verhoudingen waren veel kleiner waardoor hogere stimulus niveaus (>30 dB) nodig waren om de vooropgestelde kwaliteitscriteria te behalen. Dit was gedeeltelijk t.g.v. een ander opnameplaats, namelijk het promontorium van het slakkenhuis in primaten vs. het ronde venster in katten en chinchilla's. Een vergelijking tussen de mens en de andere diersoorten gaf aan dat de CAP-afstemming in mensen 1,6 keer scherper is dan in de kat en chinchilla en gemiddeld 1,35 keer scherper dan in de aap (makaak).

C) De geschatte neurale afstemming bij de mens, gebaseerd op de gemiddelde correctiefunctie van de andere diersoorten, toont aan dat deze scherper is dan deze gemeten van individuele zenuwvezels in chinchilla, kat en aap. Daarenboven komt de menselijke neurale afstemming, wanneer geschat op basis van de correctiefunctie van makaak, kwantitatief zeer goed overeen met de eerdere schattingen van gedragsmatige en emissiematige metingen.

In de laatste twee studies van dit proefschrift werd de neurale fasekoppeling in kat, aap en mens bestudeerd met behulp van stimulus geëvoceerde massa potentialen. Om dit te bereiken, werden verschillende technieken gecombineerd om de gesynchroniseerde fasegekoppelde neurale bijdragen te scheiden van de neurale niet-fasegekoppelde componenten (bijv. CAP) en van de niet-neurale fasegekoppelde componenten (bijv. cochleair microfonie).Voorwaartse maskering was hiervoor de cruciale techniek. Voordat de methode werd toegepast in normaal horende vrijwilligers was deze eerste geëvalueerd op de auditieve zenuwbaan en het ronde venster in katten; en daarna gevalideerd door toediening van Tetrodotoxine. Deze studies hebben geleid tot volgende bevindingen:

A) We toonden aan dat de neurale fasegekoppelde bijdrage (de neurofonie) kan worden geïsoleerd en gekwantificeerd met de hier ontwikkeld methode. Verder toonden we aan dat de neurofonie een substantieel deel uitmaakt van de massa potentiaal gemeten aan het ronde venster. Een vergelijking tussen deze neurofonie en deze rechtstreeks gemeten op de auditieve zenuwbaan toonde aan dat beide in vele opzichten zeer vergelijkbaar zijn. De enige noemenswaardige verschillen waren een typisch grotere magnitude en kleinere latentie voor de neurofonie van het ronde venster. De detecteerbare limiet van de neurofonie aan het ronde venster in katten was 4,7 kHz en dus bijna identiek is aan de ~5 kHz limiet gerapporteerd voor individuele zenuwvezels.
B) De metingen in mens en aap (makaak) vertoonden veel overeenkomsten met deze in de kat, met onder andere gelijkaardige temporele en spectrale eigenschappen. Het meest opvallende verschil in de neurofonie van primaten t.o.v. de kat was een veel kleinere magnitude (-30 dB) en signaal-tot-ruis(vloer) verhouding (-20 dB).
C) De detecteerbare frequentie limiet voor de neurofonie aan het promotorium is 3,3 kHz in de mens en 4,0 kHz in de aap (makaak). Vergelijkende data voor individuele zenuwvezels in makaak tonen een bovengrens die ong. 1,2 maal hoger is. Dit suggereert dat de bovengrens voor individuele zenuwvezels bij de mens iets hoger is dan 3.3 kHz en dus ergens ligt tussen de 3,3 en 4,0 kHz.

Algemeen Samengevat kunnen we op basis van de resultaten in dit proefschrift minimum stellen dat, ten eerste, er geen ondersteuning is voor de hypothese dat neurale fasekoppeling in mensen veel hoger zou zijn dan in de kat; ten tweede, dat onze data wel ondersteuning biedt voor de hypothese dat mensen een scherpere neurale afstemming hebben t.o.v. veelgebruikte laboratoriumdieren.

Bibliography

List of Publications

Articles in International Reviewed Journals

[1] Joris PX, Verschooten E (2013) "On the limit of neural phase-locking to fine-structure in humans", Basic Aspects of Hearing: Physiology and Perception, Eds. BCJ Moore, RD Patterson, I Winter, RP Carlyon, HE Gockel (Springer, NY), p. 101-108.

[2] Verschooten E, Desloovere C, Joris PX "Electrophysiological evidence for sharp cochlear tuning in humans", in preparation.

[3] Verschooten E, Robles L, Joris PX "Upper frequency limit of phase-locking in stimulus evoked potentials", in preparation.

[4] Verschooten E, Desloovere C, Joris PX "Upper frequency limit of phase-locking in human and monkey", in preparation.

[5] Verschooten E, Robles L, Kovačić D, Joris PX (2012) Auditory nerve frequency tuning measured with forward-masked compound action potentials. Journal of the Association for Research in Otolaryngology : JARO 13:799-817.

[6] Mc Laughlin M, Verschooten E, Joris PX (2010) "Oscillatory Dipoles As a Source of Phase Shifts in Field Potentials in the Mammalian Auditory Brainstem", J. Neurosci. 30: 13472-13487.

[7] Kuhn R, Decoutere S, Caymax M, Vleugels F, Verschooten E, Loo R (1999) "A 0.35um BiCMOS Process with Selective Epitaxial SiGe Bipolar Transistors", Solid-State Device Research Conference, Volume: 1, p. 436- 439.

Abstracts: International Conferences

[8] Eric Verschooten, Luis Robles, Christian Desloovere, Philip X. Joris, (2013) "Assessment of Neural Phase-Locking at the Round Window in Human"Assoc Res Otolaryngol Abs 36:436

[9] Goldwyn, Myles McLaughlin, Eric Verschooten, Philip Joris, John Rinzel, (2013) "A Biophysical Cable Model for MSO Neurons as Generators of the Neurophonic"Assoc Res Otolaryngol Abs 36:352

[10] E Verschooten PX Joris, (2012) "Assessment of the frequency limit of phase-locking at the round window in human" Program No. 462.01. Neuroscience Meeting Planner. New Orleans, LA: Society for Neuroscience, Online.

[11] PX Joris, E Verschooten (2012) "Upper frequency limit of phase-locking in humans: Implications from binaural hearing" Program No. 462.15. Neuroscience Meeting Planner. New Orleans, LA: Society for Neuroscience, Online.

[12] JH Goldwyn, PX Joris, M Mc Laughlin, E Verschooten, J Rinzel (2012) "Biophysically-based models of local fields and ephaptic coupling in the auditory brain stem" Program No. 364.14. Neuroscience Meeting Planner. New Orleans, LA: Society for Neuroscience, Online.

[13] Verschooten E, Desloovere C, Joris PX (2012) "Neural tuning measured with compound action potentials in normal hearing human volunteers", Assoc Res Otolaryngol. Abs 35:360

[14] Verschooten E, Robles L, Joris PX (2011) "Assessment of Phase-Locking Using Population Responses at the Auditory Nerve and Round Window", Assoc Res Otolaryngol. Abs 34:222-223

[15] Verschooten E, Robles L, Kovačić D, Joris PX (2010) "Neural tuning measured with forward-masked compound action potentials", Assoc. Res. Otolaryngol. Abs.: 25

[16] Verschooten E, Mc Laughlin M, Joris PX "Reversible inactivation of the ventral auditory brainstem by cooling", Program No. 259.10. 2009 Neuroscience Meeting Planner. Chicago, IL: Society for Neuroscience, 2009. Online.

Presentations

Conferences

[1] Eric Verschooten, Luis Robles, Philip Joris (2013) "Neural phase-locking assessed at the round window", presentation, meeting Werkgemeenschap Auditief Systeem, Eindhoven, May 3.

[2] Eric Verschooten, Christian Desloovere, Philip Joris (2011) "Neural frequency tuning in human volunteers predicted from inter-species compound action potential relations", presentation, meeting Werkgemeenschap Auditief Systeem, Rotterdam, December 20.

[3] Eric Verschooten, Luis Robles, Damir Kovačić, Philip Joris (2010) "The sharpness of compound action potential tuning curves using notched-noise forward Masking", presentation, meeting Werkgemeenschap Auditief Systeem, Leuven, May 19.

Professional Career

Eric Verschooten was born in 1972 in Duffel, Belgium and received his secondary school education in Lier at the Vrij Technisch Instituut where he studied electronics. He graduated with a master's degree of Industrial Sciences in Electronic Engineering (option telecommunication) from K.I.H. De Nayer, Mechelen, Belgium in 1995. He pursued his studies and graduated as Master of Science in Electronic Engineering (option microelectronics) from the University of Leuven, Belgium in 1998. After his studies he joined IMEC Leuven in Belgium as process engineer on CMOS and BiCMOS transistor technologies. In 2000 he changed to DVE Brussels in Belgium and worked as electronic developer/technical leader on television projects at Philips Bruges in Belgium and on LED projects at Philips Lighting Eindhoven in The Netherlands. During this period a patent was granted to him as designer (US Patent Application No 8111014; Issued on 07/02/2012). Thereafter he joined in 2006 the Association of the University of Antwerp – Artesis Hogeschool Antwerpen – in Belgium as a researcher, instructor and supervisor and collaborated with 3WIN, a spin-off of the University of Antwerp that developed the next generation hybrid cochlear implants. In 2007 he joined the Laboratory of Auditory Neurophysiology at the University of Leuven to conduct fundamental research under the supervision of Prof. Philip Joris.