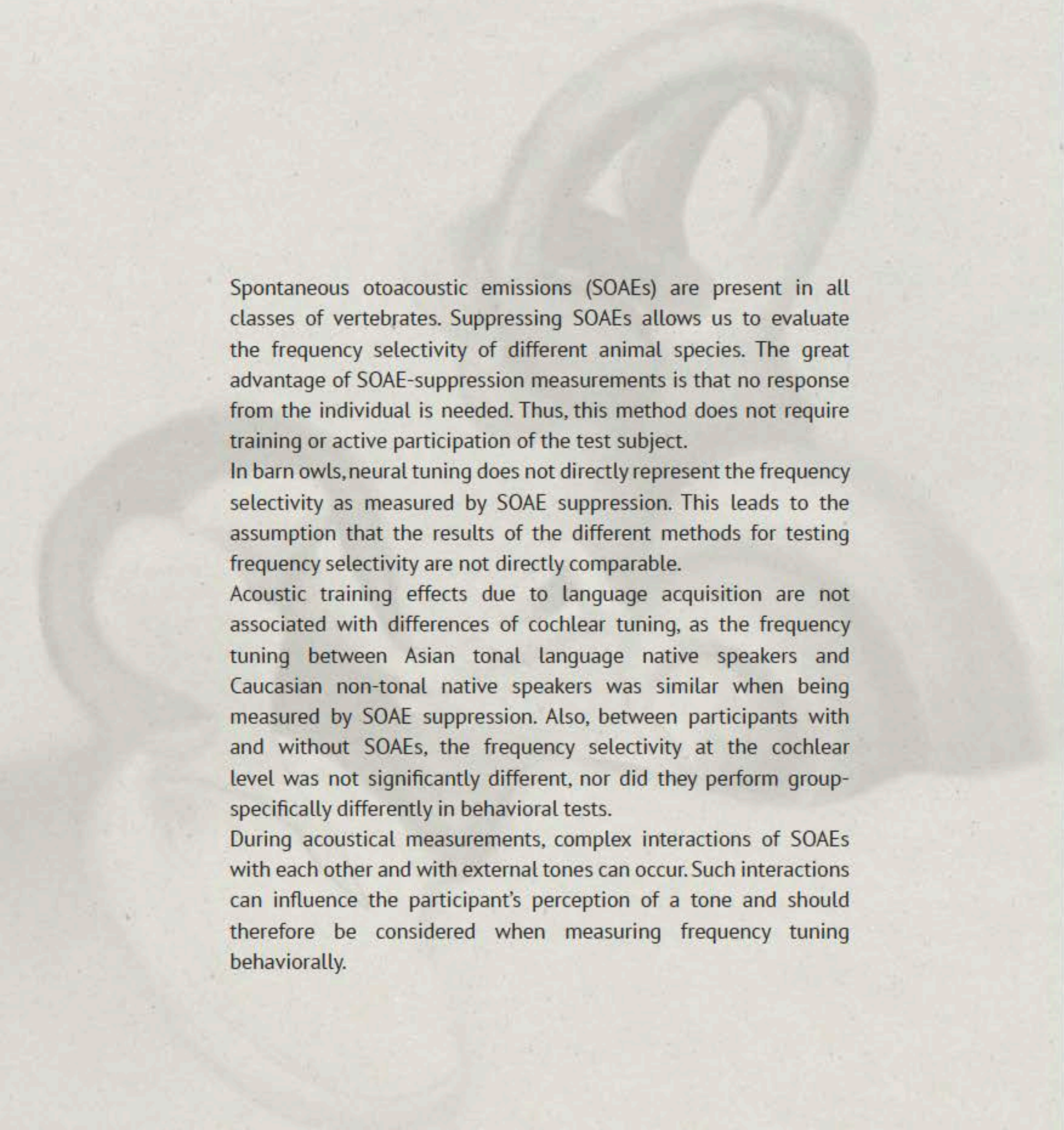


Spontaneous otoacoustic emissions (SOAEs) are present in all classes of vertebrates. Suppressing SOAEs allows us to evaluate the frequency selectivity of different animal species. The great advantage of SOAE-suppression measurements is that no response from the individual is needed. Thus, this method does not require training or active participation of the test subject.

In barn owls, neural tuning does not directly represent the frequency selectivity as measured by SOAE suppression. This leads to the assumption that the results of the different methods for testing frequency selectivity are not directly comparable.

Acoustic training effects due to language acquisition are not associated with differences of cochlear tuning, as the frequency tuning between Asian tonal language native speakers and Caucasian non-tonal native speakers was similar when being measured by SOAE suppression. Also, between participants with and without SOAEs, the frequency selectivity at the cochlear level was not significantly different, nor did they perform group-specifically differently in behavioral tests.

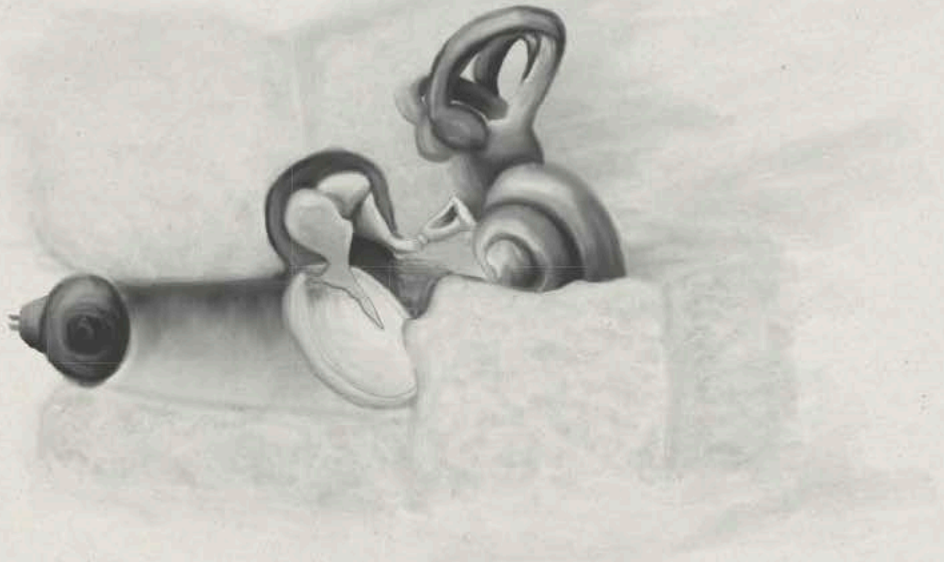
During acoustical measurements, complex interactions of SOAEs with each other and with external tones can occur. Such interactions can influence the participant's perception of a tone and should therefore be considered when measuring frequency tuning behaviorally.

Using spontaneous otoacoustic emissions to probe frequency selectivity

Sina Engler

2023

Using spontaneous otoacoustic emissions to probe frequency selectivity



Sina Engler

University of Groningen

Using spontaneous optoacoustic emissions to probe frequency selectivity

Engler, Sina

DOI:
10.33612/diss.553345668

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Engler, S. (2023). *Using spontaneous optoacoustic emissions to probe frequency selectivity*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. https://doi.org/10.33612/diss.553345668

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 20-01-2023

Using spontaneous otoacoustic emissions to probe frequency selectivity

Sina Engler

university of groningen

umcg

bcn

Huizinga Stichting

All illustrations: Sina Engler
Printing: Gildeprint Enschede, gildeprint.nl
Layout and design: Hans Schaapherder, persoonlijkproefschrift.nl

university of
groningen

Using spontaneous otoacoustic emissions to probe frequency selectivity

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Wednesday 8 March 2023 at 09.00 hours

by

Sina Engler

born on 30 March 1989
in Forst, Germany

Supervisors

Prof. P. van Dijk
Prof. N. M. Jansonius

Co-supervisor

Dr. E. de Kleine

Assessment Committee

Prof. J. C. M. Smits
Prof. J. M. C. van Dijk
Prof. J. L. Verhey

Table of Contents

Chapter 1

Introduction

Chapter 1

Preface

Hearing – a sensation that seems to easily appear to us in our daily life. But the sense of hearing is much more than just as if perceived incidentally. Hearing is the key element in human communication and plays an important role in cultural aspects as in language and music. We use acoustical information every day, while always depending on the two coiled structures in our skull, each of them about 3.5 cm long - the cochleae. The cochlea is fundamental for a highly frequency-selective acoustic perception. Frequency selectivity is the ability to discriminate between sounds with different frequencies. The focus of this thesis lays on the application of spontaneous otoacoustic emissions to measure frequency selectivity of the cochlea and thus mainly at the level of the peripheral auditory processing.

Understanding the precise mechanisms of the inner ear physiology that define the high resolution of frequency selectivity, will expand our understanding of cochlear mechanics and species-specific differences as well as the evolutionary development of the auditory system.

This first chapter introduces the general properties of acoustic cues. Further, the anatomy and physiology of the human ear will be outlined and briefly compared with the characteristics of the avian ear morphology and sound transmission. Secondly, what spontaneous otoacoustic emissions are and how they are used to test frequency selectivity of an individual will be explained. Finally, the following chapters of this thesis will be introduced.

1.1 Properties and perception of sound

Sound in air can be understood as an oscillating movement of air molecules that causes differences in density and pressure. The unit of sound pressure is pascal (Pa). The pressure variations act as traveling pressure waves that spread away from the sound source causing them. In sophisticated communication cues (as speech) such pressure variations are complex, but can be separated into their smallest components - pure tones.

Any pure tone consists of a sinusoidal waveform. Changes in increased and decreased density and pressure create the waveform of a pure tone with molecule condensation peaks and rarefaction troughs. Thus, a sound wave consists of repeating cycles of peaks and troughs. Each waveform has a specific frequency that describes how many of these waves pass a given point in one second. The frequency is consequently a measure of cycles per second, expressed in hertz (Hz). I want to emphasize here that pitch is the subjective percept of frequency. This percept correlates with the frequency as high frequencies are perceived as high pitch sounds and low frequencies as low pitch sounds.

Besides its frequency, sound is also defined by its intensity. The intensity of a sound depends on the magnitude of the molecule movements and thus the variations in pressure relative to the atmospheric pressure. A high sound intensity describes consequently a large wave amplitude. Sound intensity can be measured as sound pressure level (SPL), expressed in the unit of decibel (dB). Large sound amplitudes (intensities) are subjectively perceived as louder sounds and vice versa. Thus, the sound wave of a pure tone can be described as a sine function defined by a specific frequency and amplitude.

The human ear can detect frequencies ranging from 20 Hz to nearly 20 kHz. Sounds below the hearing range of humans are named infrasound and sounds above the frequencies of human perception are termed ultrasound. To measure the human hearing range the hearing threshold at 1 kHz was defined to be 0 dB SPL (Fletcher and Munson, 1933).

This threshold corresponds to a sound pressure level of 20 µPa. Humans are able to detect sounds of levels between 0 and 130 dB SPL. Everything above 90 dB SPL is considered to be harmful and may cause damage to the auditory system. With increasing age, the ability to detect, especially high-frequency sounds, decreases. This age-related hearing loss (presbycusis) develops progressively and is irreversible.

1.2 Peripheral auditory anatomy and physiology

The human ear can be separated into the outer, middle, and inner ear (Fig. 1.1) The outer ear consists of the pinna and the ear canal. The pinna plays an important role in directional hearing (Gardner and Gardner, 1973). The ear canal functions as a funnel that directs the sound into the ear canal towards the middle-ear. At the border between the outer and middle ear sits the tympanic membrane. A flexible structure with a wide diameter that is in contact with the smallest bones in the human body, the ossicles (malleus, incus, stapes). As the tympanic membrane moves, according to the pressure fluctuations of the sound waves that deflect it, the movement is transmitted on to the ossicles. The small ossicles transmit the movement to the oval window and this vibration causes the movement of the cochlear fluid that displaces the basilar membrane (BM) and the round window. Due to the negligible compressibility of cochlear fluid, the round window works as a compensator, moving in the opposite direction as the oval window.

As you can easily imagine, movement in the air is easier than the same movement under water. The signal transmission from air into a liquid-filled structure requires impedance matching, otherwise most of the signal would simply be reflected at the oval window. The three middle ear bones function as an impedance matching device, as the airborne signal (sound) is now a mechanical movement that needs to be passed on to the fluid-filled cochlea. The size of the stapes footplate contacting the oval window is the key to keep the signal intensity; its area is about 20 times smaller than the area of the tympanic membrane. Due to this difference in surface area the middle ear provides the necessary pressure gain to overcome the impedance mismatch from air to liquid without losing signal intensity.

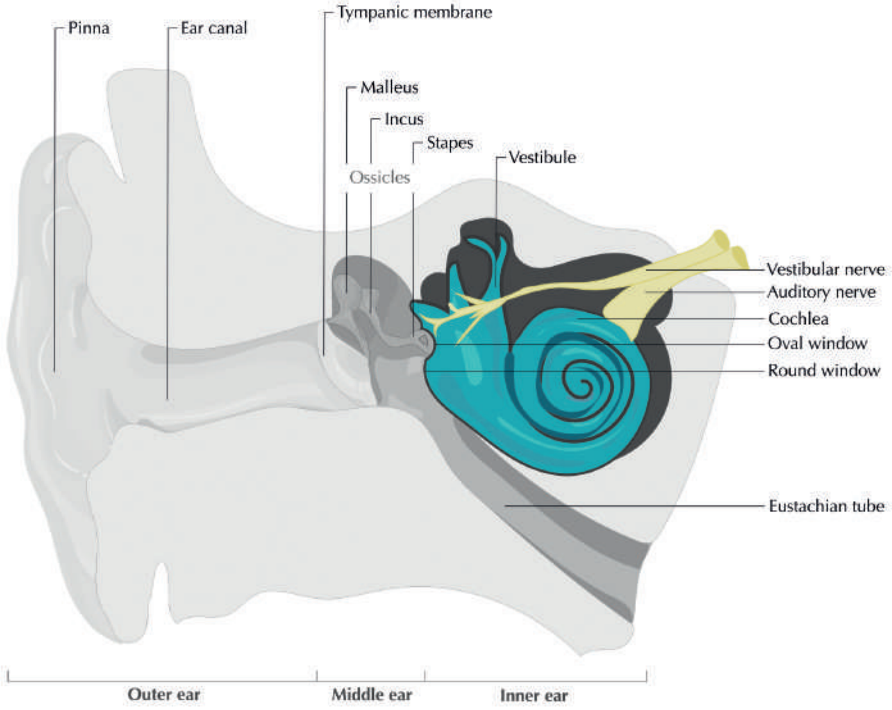


Figure 1.1. Schematic illustration of the cross section of the human ear. The peripheral auditory system consists of the outer, middle and inner ear. The cochlea is a bony structure, coiled 2.75 times that contains the sensory unit of the inner ear, the organ of Corti.

Thus, with constant force, a greater pressure is exerted onto a smaller surface. You may think of it as the force along the sole of a shoe driven by the weight of a person. Assuming the same person (constant force) is standing with a flat heel sole or a spike heel on the ground, the spike heel will have a much higher pressure on the smaller area contacting the ground, compared to the flat sole. Therefore, the latter one will be much less comfortable when standing on your foot, too. Certainly, this is a very simplified example which does not represent the complexity of realistic middle ear impedance models. However, it illustrates the benefit in reducing the area when transmitting the signal from the tympanic membrane to the stapes and ultimately to the oval window.

The shape of the cochlea is often compared to that of a snail shell, as it spirals from the base (near the stapes) to the apex upwards and inwards while coiling about 2.75 times.

Moreover, the bony structure of the cochlea also protects its inner compartments. When cross sectioning through the middle of the cochlea, the inner arrangement of the spiral becomes visible (Fig. 1.2A).

1.2.1 Organ of Corti

The inner structure of the cochlea is characterized by three fluid-filled compartments: the scala vestibuli, the scala media, and the scala tympani (Fig. 1.2A). The Reissner's membrane defines the lower border of the scala vestibuli and the upper border of the scala media. The BM defines the border between the scala media and the upper side of the scala tympani. All three compartments range through the whole length of the cochlea, from base to apex. Only at the very apical tip, the two outer compartments (scala vestibuli and scala tympani) merge at the helicotrema, where the BM is not sealed to the wall of the cochlea. At the cochlear base the scala vestibuli meets the oval window, whereas the scala tympani has contact to the round window. These two scalae share the same fluid, the perilymph, whereas the scala media contains endolymph. The positive electrical potential of the endolymph of about 80 mV relative to the perilymph is generated by potassium ions (K^+) which are actively supplied by the stria vascularis. The stria vascularis is an epithelium with blood capillaries that forms a major portion of the wall of the scala media. The potential difference between the endo- and perilymph is essential for the signal transmission of stimuli. Consequently, impairments of the stria vascularis result in hearing deficits.

The sensory unit of the cochlea is called the organ of Corti (Fig. 1.2B) which sits on the BM and is covered by the tectorial membrane (TM). The organ of Corti comprises the sensory hair cells, their supporting cells (which will be omitted from further discussion), and the innervating nerve fibers.

1.2.2 Basilar membrane and tonotopy

Frequency selectivity is fundamentally realized by the morphology of the cochlear partition. When imagining the BM stretched out, the membrane is narrow and stiff at the base and becomes increasingly wider and flaccid towards the apex. When a pure tone stimulus is presented, the fluid wave within the cochlea causes the BM to move instantaneously. The structure of the BM causes high frequencies to resonate at the base and low frequencies to resonate at the apex. Thus, the BM moves according to the presented sound, forming *traveling waves* along the membrane. The maximum wave amplitude for a given frequency occurs on a specific place on the BM. In other words, the BM movement is maximally selective for a characteristic stimulus frequency. This mechanical analysis (spatial coding) is essential to differentiate between tones, and forms the basis of the cochlea's *frequency selectivity* (or frequency tuning).

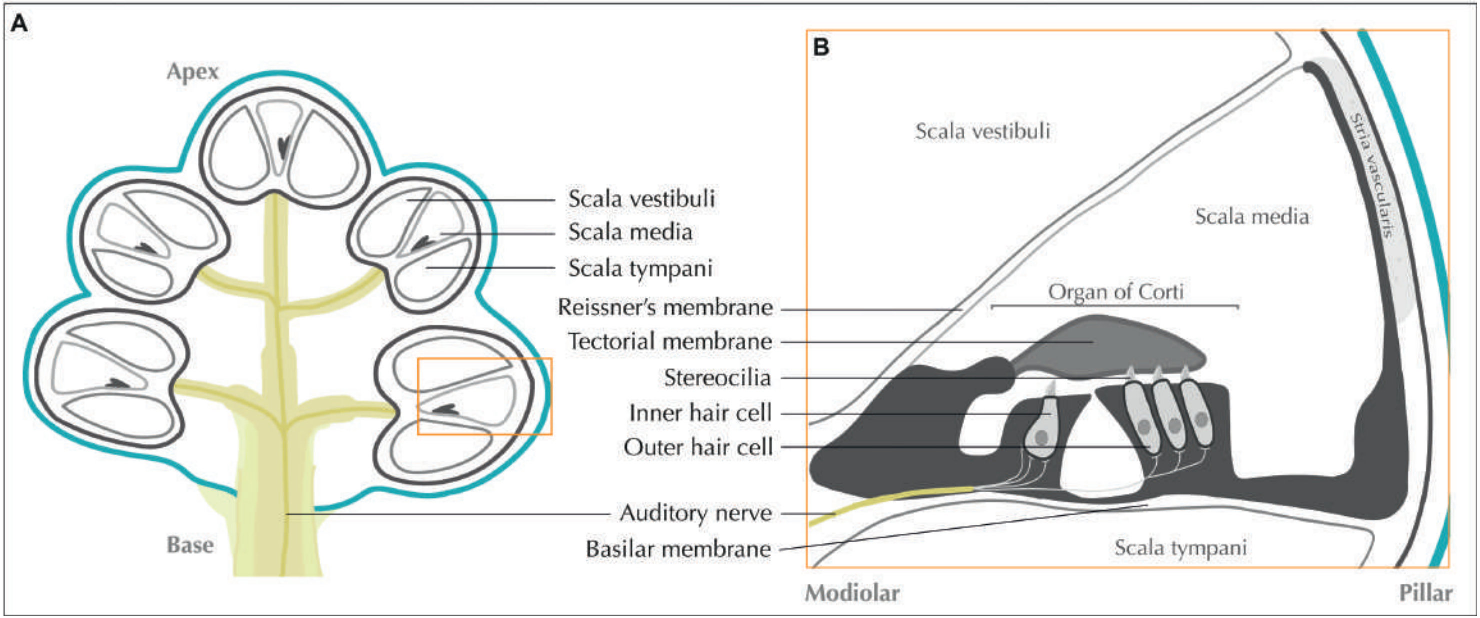


Figure 1.2. A schematic illustration of the cross-section of the cochlea. (A) The cochlear spiral is viewed from the side. The three fluid-filled compartments that spiral around the central axis can be seen. (B) A magnification of the median compartment, the scala media, reveals the organ of Corti.

The tonotopic organization is described as an array of overlapping auditory filters which are tuned to a specific frequency, their center or characteristic frequency (postulated by von Helmholtz, 1862). Von Békésy observed such place-specific BM oscillations while playing a pure tone, showing that the cochlea behaves like an array of tuned oscillators that are sensitive to a specific frequency (von Békésy, 1960). Thus, different auditory filters are excited depending on the stimulus frequency. While each filter is tuned to a characteristic frequency, the interactions of the auditory filters to their adjacent neighbors form a transmission line.

The characteristic motion of a traveling wave from the base to the apex along the BM arises because of the phase delay (Plack, 2018). It seems like the traveling wave grows in amplitude while decreasing in wavelength when moving along the BM to the specific place of maximum resonance (Fig. 1.3A).

In a fluid-filled structure as the cochlea, damping plays an important role and besides the BM mechanics, the cochlear fluids influence the frequency selectivity substantially. Not only is the BM oscillation damped, the surrounding fluid also moves. Thus, active filtering is essential to counteract the viscous damping and achieve the high resolution of frequency selectivity (postulated by Gold, 1948). Active amplification means that the output exceeds the input, resulting in a sharp (well-tuned) oscillation peak even in a fluid-filled system. Thus, frequency selectivity depends besides the physical mechanics of the BM and the cochlear fluids, also on some sort of physiological amplifier (Fig. 1.3B).

1.2.3 Hair cells

The hair cells (HCs) are based on the BM and are the sensory units in the cochlea (Fig. 1.2B). Cochlear HCs are characterized by the stereocilia (*the 'hairs'*) on their apical end. The entire hair bundle of one vertebrate HC consists of 20 to 300 stereocilia (Hudespeth, 1989; 2008). Each bundle is organized according to stereocilia length. The stereocilia towards the inner turn of the cochlea (modiolar, neural side) are the shortest, gradually increasing in height towards the outside (pillar, abneural side). The stereocilia are surrounded by endolymph while the cell body itself is placed in the perilymph.

The apical tops of the stereocilia are connected via tip-links (Pickles et al., 1984). The apical part of the tip-link connects to the next taller stereocilium and the base of a tip-link contains the mechano-electrical transduction channel. When the BM moves in response to a stimulus, the stereocilia are deflected. Stereocilia movement towards the pillar side causes the tip-links to mechanically open the transduction channel, allowing the influx of positively charged ions. Because of this mechanical action tip-links are also referred to as *gating springs*. The gating springs sense the hair bundle deflection and mechanically convert the stereocilia displacement into a force which opens the ion channels. This mechano-electrical transduction mechanism is described by a gating spring model (Howard and Hudspeth, 1988). Gating springs can adjust their stiffness to modulate the tension on the channels (Bartsch et al., 2019). How the transduction channels of the HCs are controlled by the tip-links is not further discussed here.

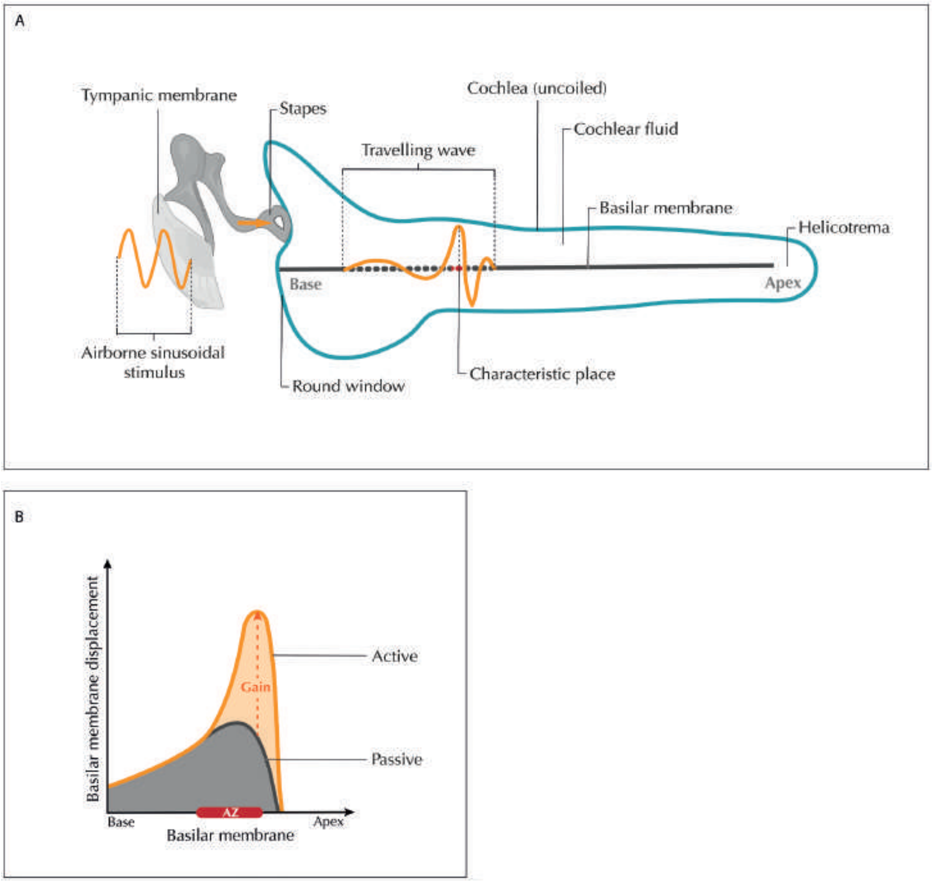


Figure 1.3. Schematic illustration of a traveling wave on the basilar membrane elicited by a pure tone. (A) The representation of the longitudinal section of an uncoiled cochlea. The sinusoidal tone deflects the tympanic membrane which then causes movement of the stapes (→) that excites the BM, resulting in a traveling wave. The highest amplitude of the traveling wave is at a specific place on the BM (•), according to the stimulus frequency. High frequencies excite the base of the BM, whereas low frequencies excite apical regions. This spatial coding of frequencies is called tonotopy. (B) Mechanical amplification of the traveling wave. The passive response (—) illustrates the traveling wave on the BM without amplification. The physiological response of the cochlea is active (—). The active gain by the outer hair cells (OHCs) changes the filtering characteristics of the BM and sharpens over a restricted range of frequencies the frequency selectivity. The peak enhancing activity of the OHCs results from an active zone (AZ) basal to the peak.

When the channels are open, potassium ions (K^+) enter the HC which causes depolarization. With the depolarization of the HC membrane, electrical gated calcium (Ca^{2+})

channels open. Subsequently, Ca^{2+} enters the cell (from the perilymph) and triggers the release of a messenger substance (neurotransmitter). HCs are in contact with sensory neurons via basal ports (synapses). The neurotransmitter is released at the base of the HC into the synaptic cleft, where synaptic terminals of the spiral ganglion neurons connect to the sensory cells. The released neurotransmitter of the HC excites the associated sensory neurons, causing the generation of action potentials. These electric potentials are transmitted to the brain, where they can be decoded and perceived as the acoustical signal.

Thus, stereocilia movements play a key role, as they regulate the mechano-transduction and thus the depolarization of the HC, which is the start of the signal cascade. The HC hyperpolarizes when the stereocilia move towards the modiolar side and the tension on the tip-links is reduced.

HCs can be divided into two types, the inner and the outer hair cells (respectively: IHCs, OHCs). Both HC types differ in their function, morphology, number, alignment, and innervation. The IHCs can be seen as the actual sensory cells that send the acoustical information as electrical signals via the auditory nerve fibers to the brain. Approximately between 3400 and 4000 IHCs line up in a single row at the modiolar side (Dallos, 1992; Hudspeth, 2014; Rask-Andersen et al., 2017). Their stereocilia are not in contact with the TM and each bundle forms a U-shaped organization.

About 12 000 OHCs form three rows at the pillar side of the cochlear turn (Dallos, 1992; Hudspeth, 2014), while each bundle of stereocilia forms a V-shape. OHCs are thought to function as the amplifiers of the ear, as they enhance the stimulus by mechanically amplifying the sound-driven vibrations. Their signal amplification (Fig. 1.3B) is essential for normal hearing. The stereocilia tips of the OHCs are in contact with the TM, which enhances the shear force at the bundles when the BM moves. Another essential difference to IHCs is that the OHC membrane contains the protein prestin which modifies the length of the cell body (Zheng et al., 2000; Dallos et al., 2008). When the OHC depolarizes, the prestin molecules shorten, which also reduces the entire length of the OHC-soma; the opposite happens during repolarization (Brownell et al., 1985; Kachar et al., 1986; Ashmore, 1987). Since this modification is induced by voltage changes it is also called *electromotility*. Thus, when the OHC ‘contracts’, the shear force between the TM and the BM increases (Ashmore et al., 2010). This somatic motility of the OHCs affects the BM motion and actively amplifies the response to a specific signal. This is how OHCs are thought to *actively amplify* (Fig. 1.3B) the motion amplitude of the BM and narrow the response peak (Plack, 2018). The active filter in the auditory system, necessary to overcome damping and allow high frequency selectivity (as postulated by Gold, 1948), is thus realized by the OHC motility. The OHCs function as a positive feedback loop (Zwicker, 1979; Dallos, 1996; Patuzzi, 1996) that connects the filter output (depolarization, shortening of the cell body) back to the input (BM movement). Active OHC motility is essential to amplify the (specific) signal by 100 to 1000 times (Ashmore, 2008) and is therefore indispensable for a high degree of frequency selectivity and sensitivity (Davis, 1983; Ruggero and Rich, 1991; Dallos, 1996; Dallos et al., 2008).

Mammalian animal models with knocked-out prestin indicated that this protein is indeed the responsible factor for the electromotility of the OHCs (Liberman et al., 2002; Dallos et al., 2008). Consequently, OHC dysfunction causes drastic losses in frequency selectivity and increases the audiometric thresholds (e.g.: Kiang et al., 1976). The active amplification reaches its effectiveness limits, with increasing signal level. At high sound levels (above 60 dB SPL) wider areas of the BM are stimulated, causing the OHCs to not selectively amplify the movement and thus the frequency selectivity decreases as well (Ashmore, 2008).

However, even though OHCs were shown to have many properties which make them the potential amplifier in the auditory system, their role as amplifiers is still under debate (van der Heijden and Vavakou, 2021). Further characteristics that support the amplifying function of the OHCs are described here in paragraph 1.5 and study results that also imply that OHCs play a crucial role in amplification mechanisms are discussed in Chapter 6.

1.3 Neural signal transmission

From the peripheral auditory system, the information is transmitted via afferent neuronal fibers (spiral ganglion neurons) to the central auditory system. More than 30,000 neurons innervate the cochlea (Otte et al., 1978). Both HC-types form synapses with different types of afferent (ascending) neurons. IHCs are innervated by type I fibers, whereas OHCs are innervated by type II fibers. Between 90 and 95% of the fibers are of the myelinated type I and the remaining 5 to 10% are of the unmyelinated type II (Spoendlin, 1985). The fibers do not only differ in the HCs they innervate, they also show differences in their spontaneous activity and their dynamic range. The latter two will not be discussed further.

Each IHC is innervated by approximately 20 type I fibers (Spoendlin, 1985; Liberman et al., 1990), while every fiber connects to one IHC (Brown, 1987; Lieberman, 1980a). This means that BM movement at each specific tonotopic place is transmitted by about 20 neurons and that every neuron is sensitive to a specific (characteristic) frequency. IHCs do not receive direct efferent (descending) feedback, instead, efferent fibers terminate on the dendrites of the afferent fibers (Raphael and Altschuler, 2003).

Each OHC is connected to a few afferent type II fibers while every fiber innervates multiple (on average up to 20) OHCs (e.g.: Berglund and Ryugo, 1987; Brown, 1987; Weisz et al., 2012). OHCs are directly innervated by efferent fibers (Raphael and Altschuler, 2003), which means that they have direct synaptic contact to afferent *and* efferent fibers, in contrast to IHCs. Thus, IHCs and OHCs differ in their afferent and efferent innervation patterns.

The neurons bunch together and form the auditory nerve while preserving the tonotopic organization: neurons of high characteristic frequencies synapse with HCs at the cochlear base and are located at the periphery of the auditory nerve. Hence, neurons of low characteristic frequencies innervate HCs at the cochlear apex and can be found in the center of the nerve bundle (Plack, 2018).

1.4 Central auditory processing

The tonotopic organization is also preserved in the central auditory system. The central auditory system reaches from the brainstem to the auditory cortex. At this central level the information from the afferent neurons is integrated and processed. The central auditory system can, however, also affect the peripheral system via efferent neuronal connections. The auditory efferent system originates from the olivary complex (Ciuman, 2010). Since efferent fibers make synaptic connections to the OHCs, this efferent feedback can directly control the OHC and consequently the BM motion. Such a central controlled '*intervention*' is thought to prevent acoustic trauma or to optimize perception.

One example of efferent influence on the auditory periphery (not involving the HCs) is the middle ear muscle (MEM) reflex. The MEM reflex is triggered by moderate to high sound levels, and can thus also be initiated while self-speaking. Two muscles form the MEM, the stapedius and the tensor tympani. A muscle contraction results in a reduced transmission of the middle ear. Furthermore, some people can control MEM-contraction, which is a clear indication of an efferent innervation (e.g.: Guinan, 1996; Liberman and Guinan 1998) and proves that the auditory periphery can be influenced by the central system.

In fact, neural plasticity seems to allow the central coordinated adaptation of peripheral auditory processing. Especially efferent neural pathways may reflect the modifications resulting from perceptual auditory learning (Plack, 2018). Therefore, language acquisition and musical training may play an important role in the plasticity of the auditory system. How such plasticity may affect peripheral frequency selectivity will be discussed in Chapter 3, where cochlear frequency selectivity of tonal and non-tonal language native speakers is compared to answer the question whether peripheral tuning differences between both groups are present. The comparability between objective and behavioral measures of frequency selectivity is further discussed in Chapter 4, aiming to extend our understanding of the comparability between different measurements of frequency selectivity.

1.5 Spontaneous otoacoustic emissions (SOAEs)

For a long time, it was thought that the transduction system of the ear functions in one direction – sound enters the ear, passes the earlier described levels of processing and is finally perceived as an auditory sensation. However, our ears do not exclusively function as microphones, they can also be speakers. In other words, human ears can not only detect the vibration of air molecules, they can also produce them (Kemp, 1978, 1979a; Probst et al., 1991). When placing a small microphone into the ear canal, tonal signals, called otoacoustic emissions (OAEs), can be recorded in the majority of healthy ears. The groundbreaking discovery of OAEs by Kemp (1978) changed the view on the functionality of the inner ear dramatically. Initially OAEs were termed '*Kemp-echoes*', as they were recorded after presenting a click stimulus. The nonlinear properties of such '*echoes*' lead to the conclusion that they are in fact not simply the reflection of the original stimulus. Not only is the delay between the click stimulus and the recorded response too long, but the response intensity does not increase linearly with the intensity of the stimulus either. Thus, the terminology of (click evoked) OAE was born.

OAEs can also be recorded in the absence of any external stimulation (e.g.: Zurek, 1981). Such emissions are called spontaneous otoacoustic emissions (SOAEs). This type of emission is generated by spontaneous OHC activity that continuously creates wavelike BM movements, which sets the ossicles and the tympanic membrane in motion. Consequently, the motion of the tympanic membrane moves the air in the ear canal. The oscillating air molecules can be recorded as a signal generated by the inner ear – the SOAE. Thus, the human ear may also work in reverse, namely as a speaker.

As previously mentioned, Gold (1948) hypothesized that active filtering is necessary for a high frequency resolution (Fig. 1.3B). Moreover, he predicted the existence of OAEs as the result of an inner ear amplifier that becomes unstable. An active filter feeds the output of the sensor back to the input. Such a filter can become unstable and self-oscillating, when the feed-back energy is too large. Due to the nature of amplifiers, the positive feedback loop may cause the unstable filter to oscillate spontaneously. The electromotility of OHCs and their efferent innervation qualifies them as the amplifier of the inner ear (Dallos, 1996; Ashmore, 2008). Therefore, SOAEs can be understood as the product of self-sustaining OHC-oscillation. Besides the amplification and the selective response to stimuli, the presence of SOAEs (Dallos, 1992) and the nonlinearity of the system defined by suppression and distortion products are characteristic implementations of an auditory filter with positive-feedback (Hudspeth 2008; Robels and Ruggero, 2011).

SOAEs are constant pure tone-like signals that show only small fluctuations in frequency and level (van Dijk and Wit, 1990). The presence of SOAEs is common, as the majority of humans with healthy ears emit them (Talmadge et al., 1993: Penner and Zhang, 1997; Pasanen and McFadden, 2000). SOAEs are usually recorded between 1 and 2 kHz, which corresponds to the most efficient frequency range that is transmitted in reverse direction through the middle ear (Zurek, 1981; Probst et al., 1991). The presence of SOAEs in normal hearing ears, more strictly speaking, in regions of normal audiometric thresholds, is strong evidence that they are related to physiological inner ear mechanics. However, SOAEs are not recorded in all normal hearing ears. They may also be considered as a by-product of active amplification related to the reflections within the cochlea. In fact, the SOAEs themselves are surely not essential for normal hearing. Still, the observation that SOAEs are related to audiometric threshold minima (Wilson, 1980; Long and Tubis, 1988a) indicates that they are also related to a physiological hearing mechanism. Details on interactions between SOAEs and psychoacoustic measures are discussed in Chapter 4.

1.5.1 Using SOAEs to probe for frequency selectivity

As SOAEs result from OHC activity, they can be seen as a window to evaluate inner ear physiology non-invasively. Remember that the BM consists of an array of auditory band-pass filters, where each filter reflects the frequency selectivity of a characteristic place (Fig. 1.3A). Every filter passes only a narrow band of frequency components. OHC responses are tuned to a specific frequency due to the tonotopic organization of the BM on which they are positioned. An SOAE generated by OHCs has a frequency that relates to the tuning frequency of the corresponding place on the BM. Consequently, SOAEs are signals, predominantly generated by OHC activity in the center of a specific auditory filter. OHC activity can be affected by frequencies within the same filter.

Frequency tuning (of a specific filter) can be measured by evaluating the frequency range of external tones that influences the SOAE. External tones can suppress SOAEs, this effect is frequency-dependent and describes the nonlinear interaction between the external suppressor and the OHC activity on the BM. The suppressive effect of external stimuli on a specific SOAE is plotted as a suppression tuning curve (STC). Thus, STCs allow the non-invasive and objective evaluation of cochlear frequency selectivity. Moreover, STCs of SOAEs allow intra- and interspecies comparisons of inner ear mechanisms, which are treated in Chapters 2, 3, and 4.

1.6 SOAEs in the barn owl

SOAEs can be found in many different species, which evidently have different ear morphologies. The fact that SOAEs are (besides in mammals) also present in birds, reptiles, and amphibians indicates that emissions are originating from fundamental inner ear characteristics shared among vertebrates. Comparing the anatomy of a human ear to that of a bird reveals major differences (see Table 1.1). The most obvious one is that birds (like all non-mammalian vertebrates) have no pinna. Other anatomical differences to mammals can be found in the middle and inner ear (Fig. 1.4). Comparisons between the frequency selectivity of humans to other species benefits the exploration of evolutionary development and the different physiological mechanisms of the auditory system.

The barn owl (*Tyto alba*) represents an appropriate model organism to study frequency selectivity in non-mammals. Not only is their hearing range very similar to that of humans (Konishi, 1973), they also have SOAEs (van Dijk et al., 1996; Taschenberger and Manley, 1997). The avian cochlear duct is slightly curved and about 11-12 mm long (Smith et al., 1985; Fisher et al., 1988). The cochlear duct contains the BM (Fig. 1.4) that is also tonotopically organized, coding high frequencies at the base. The BM is also narrower and thicker at the base and has a wider and thinner distal end (Smith et al., 1985). Moreover, barn owls do not suffer from age-related hearing loss (Krumm et al., 2017). Especially the latter has placed them in the focus of hearing research.

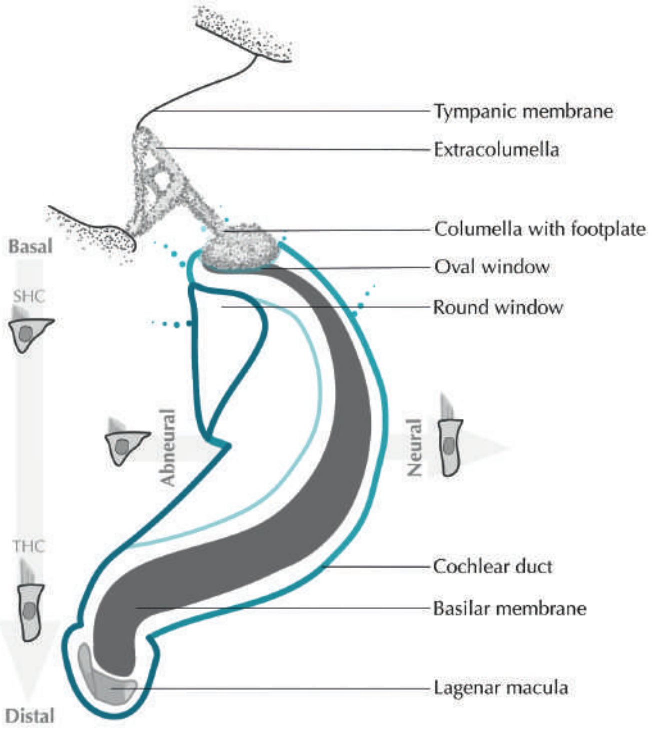


Figure 1.4. A schematic illustration of the middle and inner ear (cochlear duct) of a barn owl. The avian middle ear consists of the tympanic membrane that is in contact with cartilaginous projections of the one middle ear bone, the columella. The movement of the tympanic membrane is passed on in a piston-like movement of the columella to the oval window of the cochlear duct. The cochlear duct of the barn owl is curved and twisted and tonotopically organized. The tall and short hair cells (THC, SHC, respectively) are distributed over the entire basilar membrane and covered by the tectorial membrane, which is for illustrative reasons not shown. Hair cell types change gradually along and across the avian basilar papilla.

The answer to their *'ageless'* ears sits in their ability to regenerate HCs. Damaged avian HCs are able to be renewed, which allows birds to maintain good hearing thresholds during their entire lifespan (Corwin and Cotanche, 1988; Ryals and Rubel., 1988; Smolders, 1999). Even after severe local events such as noise injury or ototoxic damage, avian HCs can recover to nearly full functionality (e.g.: Ryals et al., 2013).

Also in contrast to mammals, the avian HCs cover the entire surface of the BM and are not aligned in rows. HCs of the barn owl can also be divided into two groups, the tall HCs (THCs) and the short HCs (SHCs). THCs are taller than wide and are characterized by long hair bundles with relatively few stereocilia (Gleich and Manley, 2000). They are located at the neural side of the basilar papilla and decrease in predominance towards the base (Fischer, 1992). About 20 to 25% of the HCs are SHCs (Köppl, 2015), characteristically with a soma that is wider than tall. This HC type has shorter hair bundles with numerous stereocilia. They are located at the abneural side of the basilar papilla. The number of SHCs is dense at the base and decreases towards the apex (Hirokawa, 1978). This cell differentiation is, however, rather gradual across the width of the basilar papilla (Fischer et al., 1988). Decisive for their clear assignment in either THC or SHC is their neural innervation. In contrast to THCs, SHCs do not receive afferent innervation (Fischer, 1992), indicating that they are clearly not the primary receptor cell. Indirect efferent innervations are not present in the avian inner ear.

HC bundle orientation changes abruptly across the width of the papilla, realized by a HC-rotation up to 90° towards the distal end (Fischer et al., 1988). The number of stereocilia in the HC bundle may also mechanically tune the frequency response (Köppl, 2015). Thus, traveling waves may be rather complex in the avian ear (Tilney et al., 1987).

Even though the precise mechanism of amplification remains unknown, the presence of SOAEs can be seen as evidence for an active amplification within the barn owl's inner ear. Due to the similarities between mammalian OHCs and avian SHCs, it is assumed that the latter play a role in the amplification process and SOAE generation in the barn owl. Avian HCs lack somatic electromotility (barn owl: Köppl et al., 2004; chicken: He et al., 2003). Thus, another mechanical amplification mechanism must be present in order to generate SOAEs. Avian HCs are electrically tuned to characteristic frequencies (Tan et al., 2013).

The density of hair cells is about 10-fold larger (Sul and Iwasa, 2009) and numbers of stereocilia are greater compared to the mammalian inner ear (Köppl et al., 2009). Thus, HC bundle motility may be the underlying mechanism of avian counteracting to damping and emission generation (Chen et al., 2001; Xia et al., 2016). However, the details of the operating principles of this active amplifier remains an open question. Frequency selectivity of the barn owl can be measured by STCs of SOAEs. Details on this measurement compared to other measures of frequency selectivity and to other species will be discussed in Chapter 2.

Table 1.1: Comparison between the human and the barn owl auditory perception and inner ear morphology.

Structure	Human		Barn owl	
Hearing range	20 Hz – 20 kHz		100 Hz – 13 kHz	
Inner ear	Cochlea, coiled, 2.75 turns		Cochlear duct, curved	
Basilar membrane	tonotopically organized uncoiled length: ± 34 mm		tonotopically organized length: ± 12 mm	
	Structure	**Sensitivity**	**Structure**	**Sensitivity**
• basal	stiff, thick, narrow	high frequencies	thick, narrow	high frequencies
• apical/distal	floppy, thin and wide	low frequencies	thin, wide	low frequencies
Covered by TM*	yes		yes	
Unit of the sensory organ	organ of Corti		basilar papilla	
Hair cells	15 000 HCs, two distinct types		16 000 HCs, gradually change from one type to another, clear differentiation in neural innervation	
Bundle orientation	aligned in rows		systematic variation along width and length	
Hair cell types	Inner Hair Cell (IHC)	Outer Hair Cell (OHC)	Tall Hair Cell (THC)	Short Hair Cell (SHC)
Function	sensor	amplifier	sensor	amplifier
Location	modiolar	pillar	Neural, mainly distal	Abneural, mainly basal
Hair cell innervation	mainly afferent, indirectly efferent	mainly efferent, also afferent	afferent & efferent	efferent
Contact to TM*	no	yes	yes, embedded	yes, embedded
SOAE* generator	OHCs electromotility		SHCs hair bundles[!]	

*SOAE=Spontaneous otoacoustic emission; TM=Tectorial membrane
[!] The details on SOAE generation in the barn owl are not known.

It can be summarized that mammalian and avian ears are ‘*homolog in their structures, but not in their specializations*’ (Köppl, 2011). The inner ear of barn owls is one of the longest among all birds, but in relation to similar sized mammals, shorter. The tuning of the BM may be the key to their perception of high frequencies (Quine and Konishi, 1974), which is exceptional among birds. The HCs of mammals and birds have a similar structure and can be divided into two main cell types that share several features. Mammalian IHCs and avian THCs are located at the neural side of the membrane, whereas OHCs and SHCs are located at the abneural side. IHCs and THCs receive much more afferent innervation compared to OHCs and SHCs, which are much more innervated by efferent nerve fibers.

Undoubtedly, the physiology of the transmission pathway is the key feature, essential for perceiving the acoustic stimuli adequately. Consequently, signals that are not transmitted at any of the stages will be irrecoverably lost and thus not available as information in later processing steps. Cochlear frequency selectivity is essential to analyze the spectral composition of a waveform. Since environmental acoustical information is usually complex, it is essential to determine the frequency composition of complex stimuli. The tonotopic organization of the BM separates the complex sound into its pure tone components, a process similar to Fourier analysis.

1.7 Aim and outline of this thesis

The overarching aim of this dissertation is the examination of frequency tuning across different species and individuals. The presented results also provide insights in the comparability of different measurements of frequency selectivity.

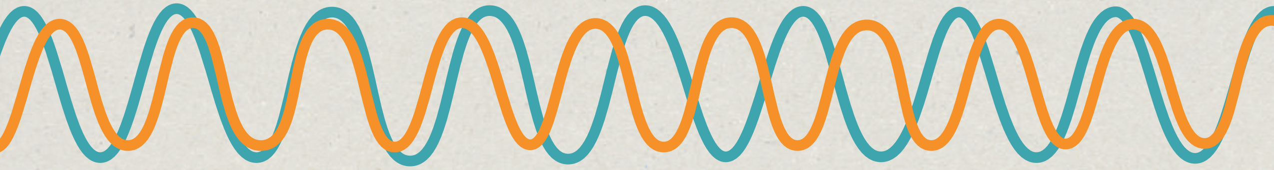
In **Chapter 2** STCs of SOAEs were measured in the barn owl and compared to other measures of frequency selectivity, to evaluate differences of different measurement methods that include different levels of auditory processing.

In **Chapter 3** STCs of two different native language groups (tonal and non-tonal) were compared to answer the question whether behavioral differences in auditory task performance can be based on differences in cochlear tuning.

In **Chapter 4** STCs and behavioral responses to tones (via psychoacoustic tuning curves, PTCs) were measured, investigating the relationship between both methods of evaluating frequency selectivity, aiming to extend our understanding of the comparability between different measurements of frequency selectivity.

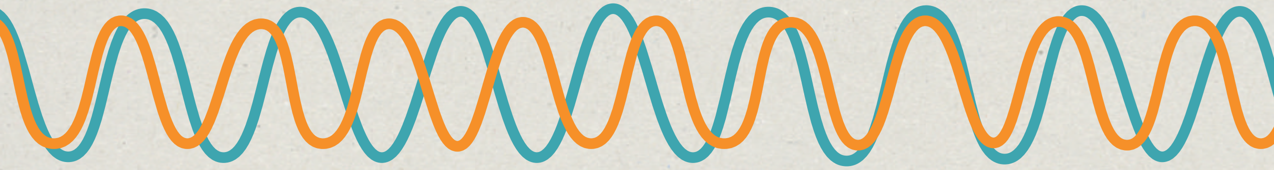
In **Chapter 5** SOAE interactions were reviewed, since SOAEs can not only be influenced by external tones (as in STC measurements), but also by other SOAEs. It is described how such interactions may influence acoustical measurements, such as pure tone audiometry.

Chapter 6 provides an overall discussion based on the study results and outlines perspectives for future research.



Chapter 2

Suppression tuning of spontaneous otoacoustic emissions in the barn owl *(Tyto alba)*



S. Engler, C. Köppl, G.A. Manley, E. de Kleine, P. van Dijk

This chapter has been published as Engler et al., 2020. Hearing Research 385, 1-9.

Abstract

Spontaneous otoacoustic emissions (SOAEs) have been observed in a variety of different vertebrates, including humans and barn owls (*Tyto alba*). The underlying mechanisms producing the SOAEs and the meaning of their characteristics regarding the frequency selectivity of an individual and species are, however, still under debate. In the present study, we measured SOAE spectra in lightly anesthetized barn owls and suppressed their amplitudes by presenting pure tones at different frequencies and sound levels. Suppression effects were quantified by deriving suppression tuning curves (STCs) with a criterion of 2 dB suppression. SOAEs were found in 100% of ears (n = 14), with an average of 12.7 SOAEs per ear. Across the whole SOAE frequency range of 3.4-10.2 kHz, the distances between neighbouring SOAEs were relatively uniform, with a median distance of 430 Hz. The majority (87.6%) of SOAEs were recorded at frequencies that fall within the barn owl's auditory fovea (5-10 kHz). The STCs were V-shaped and sharply tuned, similar to STCs from humans and other species. Between 5 and 10 kHz, the median Q_{10dB} value of STC was 4.87 and was thus lower than that of owl single-unit neural data. There was no evidence for secondary STC side-lobes, as seen in humans. The best thresholds of the STCs varied from 7.0 to 57.5 dB SPL and correlated with SOAE level, such that smaller SOAEs tended to require a higher sound level to be suppressed. While similar, the frequency-threshold curves of auditory-nerve fibers and STCs of SOAEs differ in some respects in their tuning characteristics indicating that SOAE suppression tuning in the barn owl may not directly reflect neural tuning in primary auditory nerve fibers.

2.1 Introduction

Spontaneous otoacoustic emissions (SOAEs) are sounds that are emitted by the inner ear in the absence of any stimulation. They can be recorded using a sensitive microphone in the ear canal. SOAEs appear as amplitude-stabilized signals and evidence suggests that they reflect properties of hair cells (Brownell, 1990; Manley, 2000; Kemp, 2002). Only about 60-70 percent of young, normal-hearing humans have recordable SOAEs (Talmadge et al., 1993), an indication that SOAEs are not essential for sensitive hearing in humans. Similarly, SOAEs are not shown by most laboratory animals, although their hearing sensitivity is normal. It is not yet clear why most mammalian species that were studied do not have detectable SOAEs.

Despite great variation of the inner ear anatomy, SOAEs have been described from all land vertebrate classes (e.g.: mammals: Kemp, 1979a; Ohyama et al., 1991; Talmadge et al., 1993, birds: Manley and Taschenberger, 1993; Taschenberger and Manley, 1997, lizards: Köppl and Manley, 1993; Manley, 2000, 2001, 2004, and amphibians: Palmer and Wilson, 1982; van Dijk and Manley, 2001). SOAEs share characteristics across species (Köppl, 1995; Bergevin et al., 2015), suggesting that they represent a fundamental inner ear characteristic (Bergevin et al., 2015; Manley, 2000, 2001). In lizard species, the characteristic and selective effects of suppressive tones, which enable building suppression tuning curves (STCs), show remarkable resemblances to the excitatory threshold tuning curves of single, auditory-nerve fibers (Manley and Köppl, 2008). Even though otoacoustic emissions were initially described 40 years ago (Kemp, 1979a), details regarding their origin and their significance for inner-ear function remain unexplained.

The fact that avian hair cells are able to regenerate and maintain their functionality (Langemann et al., 1999; Smolders, 1999; Ryals et al., 2013; Krumm et al., 2017) has placed birds in the focus of hearing research. Previous behavioral studies showed that starlings, *Sturnus vulgaris*, and barn owls, *Tyto alba*, do not develop presbycusis during their lifetime (Langemann et al., 1999; Krumm et al., 2017). Moreover, the avian basilar papilla is homologous to the mammalian cochlea (Manley and Köppl, 1998; Köppl, 2011; Manley, 2000, 2017) and the hearing range of barn owls covers frequencies from below 500 Hz to above 10 kHz and is thus very similar to the human range of acoustic perception (Konishi, 1973). Behavioral tests also showed that birds and mammals perform similarly when discriminating frequency or level (Dooling, 1982; review: Köppl, 2015).

Avian hearing organs have two types of hair cells that grade into each other. Of these, the short hair cells, that are defined by their lack of an afferent innervation (Fischer, 1992; Manley and Gleich, 1992; Köppl, 2011), show functional similarities to mammalian outer hair cells (Beurg et al., 2013) and may be involved in active amplification (Manley and van Dijk, 2008). Despite characteristic differences in the details of their ear morphologies, SOAE suppression has been demonstrated in both birds and mammals and thus allows the intra- and interspecific evaluation and comparison of frequency tuning.

Understanding the SOAE properties of barn owls might help elucidate their source and contribute to our general understanding of frequency selectivity.

The barn owl represents a highly specialized species and is established as a model organism for hearing research. By relying on acoustic cues, this animal can localize and catch its prey with high precision even in complete darkness (Payne, 1971; Konishi, 1973). Compared to other bird species, barn owls perceive higher frequency sounds (Konishi, 1973; Dyson et al., 1998; Krumm et al., 2017) and, due to the effects of the facial ruff, at lower sound pressure levels (review: Köppl, 2015). Moreover, the inner ear of the barn owl is complex and large, being 12 mm long (Fischer et al., 1988). In most birds, such as pigeons (Smolders et al., 1995) or chickens (Fischer, 1992), the basilar papillae are only approximately 5 mm long. The auditory sensitivity range of the barn owl ear covers about 5 octaves. Extraordinarily, the barn owl cochlea has an auditory fovea in which the highest-frequency octave (above 5 kHz) occupies half of the entire papilla (Köppl et al., 1993). Barn owls also perform remarkably fast temporal processing, with neuronal phase locking up to 10 kHz, i.e. more than an octave above the frequency ranges of phase locking shown in any other species (Köppl, 1997b).

To date, the barn owl is the only bird species in which SOAEs have been detected. Comparisons between mammalian and non-mammalian SOAEs reveal profound similarities, even though the anatomical properties of the inner ears differ significantly (Manley, 2001; Bergevin et al., 2008, 2015). Although a previous study demonstrated the existence and basic properties of SOAEs in barn owls (Taschenberger and Manley, 1997), the sample was limited due to the relatively poor sensitivity of the recording systems at that time.

Suppression of SOAEs by external tonal stimuli has been explored in several species and provides a non-invasive measure of inner-ear frequency selectivity (barn owl: Taschenberger and Manley, 1997, bobtail lizard: Köppl and Manley, 1994, Macaque: Martin et al., 1988, human: Zizz and Glattke, 1988; Manley and van Dijk, 2016). Moreover, it provides insight into inner ear mechanics, and in humans has been suggested to probe standing waves in the inner ear (Manley and van Dijk, 2016; Epp et al., 2018). In this respect it is not important whether the loss of amplitude in the presence of added tones is due to true suppression or to entrainment by the external tone. In this report, we use the term "suppression tuning".

Using a more sensitive and partly automated data acquisition system in this study as compared to the previous report (Taschenberger and Manley, 1997), we obtained a larger SOAE sample and compare details of STCs of barn owls to neuronal tuning curves from nerve fiber recordings in the same species (Köppl, 1997a, b, and unpublished results).

2.2 Material and methods

2.2.1 Animals

The measurements were carried out on seven adult barn owls (*Tyto alba*), aged between 1.5 and nearly 5 years, from the breeding colony of the Carl von Ossietzky University Oldenburg, Germany. The protocol was approved by the relevant government agency (LAVES, Oldenburg, Germany; permit number 33.9-42502-04-13/ 1182). Animals were lightly anesthetized with a combination of ketamine and xylazine to prevent movement during the measurements. They were deprived of food 12 h previously and the initial intramuscular (i.m.) injections were given immediately after capture, to minimize stress levels during the entire procedure. Initial doses were 3 mg/kg xylazine (2%, Medistar, Serumwerk Bernburg AG), and 10 mg/kg ketamine (10%, Bela-pharm GmbH & Co. KG). Light anesthesia was maintained with i.m. injections of maximally half of the initial doses every 30-100 min. The owls were placed in a double-walled, sound-attenuating chamber (Industrial Acoustics Company, Niederkrüchten, Germany) during the entire measurement. To maintain the animal's temperature between 39 and 40 °C, the body was wrapped in a feedback-controlled heating blanket connected to a rectal thermometer (Harvard Apparatus, Holliston, Massachusetts, USA). Other vital parameters, such as breathing and the electrocardiogram, were recorded via needle electrodes in muscles of a wing and the contralateral leg, and monitored using an oscilloscope and auditory monitor outside the chamber. The animals breathed unaided. The beak was fixed in a custom-made holder that maintained the position of the head during the measurements. Since middle-ear pressure in birds may fall to unnatural values under anesthesia (review: Larsen et al., 2016), the middle ear was vented via a 19G hypodermic needle set in the middle ear cavity on one side. This vent was maintained through the entire experiment. At the conclusion of the measurements, the cannula was removed and the skin incision sutured. The owl then received an i.m. injection of 0.02 ml meloxicam (2 mg/ml, "Metacam", Boehringer, Ingelheim) as an analgesic and anti-inflammatory agent for the recovery phase.

2.2.2 Recording procedure

Both ears of each owl were examined for the presence of SOAEs. The recording procedure encompassed three main steps: (1) A recording of the sound field in the ear canal without external stimuli (2-min of recording for five ears; 5-min in nine ears). (2) The suppression measurement, during which the SOAE signal was recorded while tones over a large number of levels and frequencies were presented in quasi-random sequence. The duration of this measurement was approximately 35 min and depended on the number of stimuli presented. (3) A further SOAE recording in quiet of 2 min (equivalent to step 1), to record reference values for the SOAEs and evaluate possible shifts.

An Etymotic ER-10C microphone-speaker system (Etymotic Research, Inc., Elk Grove Village, IL, USA) with a soft foam ear plug was placed at the entrance to the external ear canal, thus occluding it. The output of the microphone was amplified by 20 dB using an Etymotic ER-10C DPOAE probe driver-preamplifier (except for one individual, where a 40 dB amplification was used). To monitor the SOAE, the amplified signal was fed into a spectrum analyzer (Stanford Research System, model SR 760), covering a frequency range from 0 kHz to 16 kHz. An Audiofire ESI U 24 XL AD/DA converter (ESI Audiotechnik GmbH, Leonberg, Germany) was used to record the microphone signal on a computer disk and to generate stimuli. This converter was controlled by custom routines developed with Matlab software (2016a, MathWorks Inc., Natick, MA, USA). The AD and DA conversion were performed at 24-bit resolution and a 48 kHz sampling rate.

SOAEs were identified as peaks exceeding the noise floor and that in the averaged spectrum were suppressible by external tones. Moreover, SOAEs were individual for each ear and identifiable in both baseline measurements (step 1 and 3, described above). Small frequency components that were not amenable to the Lorentzian curve fit (van Dijk and Wit, 1990) were excluded from further analysis. In our study, the SOAE level is defined by the area under the emission peak. This method allows a precise and robust measure of emission levels, especially if the peak does not fall within one resolution bin. For the subset included in the STC analysis, we further required that the SOAE was suppressed by at least 2 dB by external tones of amplitudes lower than 80 dB SPL. The initial emission recording (step 1) was used to define the SOAE frequencies (f_{SOAE}) and levels. The average frequency of each SOAE in both unsuppressed recordings (step 1 and 3) was used to define the average frequency of the emission ($f_{average}$) used in the suppression analysis.

2.2.3 Stimulus presentation

In order to investigate suppression of SOAEs, brief stimulus tones were presented over a wide range of frequencies and levels. The duration of each tone was 1.2 s, including a 10 ms cosine rise/ fall time. SOAE recording started 150 ms prior to the tone onset and ended 150 ms after tone offset. Thus, for each stimulus tone, a segment of 1.5 s of the microphone signal was recorded and stored for later analysis. In one individual, the tone duration was 2.4 s. The stimulus frequencies were chosen to generously cover the range in which SOAEs were detected. In most cases, the suppression frequency varied from 4 to 16 kHz in 1/24 octave steps. In one individual, the step size was 1/16 octave.

The stimulus levels varied between presented frequencies and ears. The widest range was -13 to 81.2 dB SPL in 4 dB steps. In a typical case, with 49 frequencies between 4 and 16 kHz and 22 levels, the total number of stimuli was 1078. The sound pressure levels (SPLs) of the stimuli were roughly equalized according to the frequency response recorded using a Brüel & Kjaer system (type 4136) in a custom-build coupler that mimicked the acoustics of the barn-owl ear canal. Final SPLs were post-hoc corrected using the Etymotic ER-10C readings of actual stimulus levels in the individual ear canal, using a single sensitivity factor for the ER-10C.

2.2.4 Data analysis

From the microphone recording of a single tone presentation, the effect of that tone on each of the SOAE spectral peaks could be obtained. For each SOAE frequency ($f_{average}$) of interest, the following analysis was carried out.

As described above, for each stimulus tone, a recording of 1.5 s was stored: 0.15 s without stimulus, then 1.2 s with stimulus, followed by 0.15 s without stimulus. The center 1 s of this recording was evaluated. Note that the stimulus tone was on during this entire 1-s interval. The purpose of the subsequent analysis was to determine the amplitude of the SOAE of interest in the presence of the tonal stimulus.

First, a tonal signal with a frequency equal to the stimulus plus two higher harmonics was fitted to the time-domain of the recorded signal. The resulting fit was subtracted from the recorded signal. This provided a residual that included the SOAEs from the barn owl ear, but excluded the stimulus tones and its harmonics. Second, the SOAE frequency of interest was isolated by application of a zero-phase band-pass filter with an amplitude response determined by the average ($f_{average}$) and the width of the filter (Δf):

$$A(f) = \left[1 + \frac{(2[f - f_{average}])^{8}}{\Delta f}\right]^{-\frac{1}{2}}$$

The center frequency of the filter was placed at the unsuppressed $f_{average}$ and the width of the filter set to 400 Hz.

The Hilbert phase of the filtered signal was then used to compute the average of the actual SOAE frequency during the 1-s segment. Thirdly, the filter procedure was repeated, but with a filter center frequency that now equaled this computed SOAE frequency, and the filter width was narrowed to 200 Hz. Finally, from the resulting filtered signal, the SOAE level was obtained as the averaged Hilbert envelope.

As described above, the $f_{average}$ was used as the center frequency of the initial filter during the suppression analysis. Whenever the emission frequencies of the initial (step 1) and final recording (step 3) drifted by ≥ 200 Hz, this particular SOAE was excluded from the analysis (in total 9.6% of all SOAEs), since the SOAE signal would potentially drift out of the analysis filter and would not be reliably tracked.

By repeating this procedure for each of the stimulus presentations, a full frequency matrix of SOAE amplitudes was obtained. Each matrix element contained the SOAE amplitude for a specific stimulus amplitude and -frequency. This procedure was only able to reliably identify and isolate SOAEs that were more than about ±100 Hz away from a stimulus tone; for stimulus tones closer than this 200 Hz window, we were unable to assess SOAE suppression. For every stimulus frequency, the tone level at which the emission reached 2 dB attenuation was calculated. A 3-point moving average along the level and frequency dimensions was applied to create smoothed matrices. Such a data set was obtained for each $f_{average}$, whenever 2 dB attenuation was reached the smoothed amplitude matrix was computed by linear interpolation between successive tone levels. The results were subsequently combined for various frequencies to calculate STCs. Thus 2 dB STC are characterized by all relevant suppressor-tone frequencies and -levels. The lowest suppression tone level is referred to as the threshold, with a corresponding tip-frequency (f_{tip}) of the tuning curve. According to custom, the Q_{10dB} value, which describes the tuning selectivity was calculated as:

$$Q_{10dB} = \frac{f_{tip}}{\Delta f_{10dB}}$$

Where f_{tip} denotes the STC-tip frequency and $\Delta f_{10\ dB}$ the width of the STC at 10 dB above the tip level.

The slopes for the lower and the higher frequency flanks of each STC were evaluated. According to f_{tip}, and to enable direct comparisons with previous work (Taschenberger and Manley, 1997), two levels 3 dB (L_1) and 23 dB (L_2) above the tuning curve threshold and the corresponding frequencies (f_1 and f_2) were calculated by using an interpolation routine. For each STC, the slopes of the two flanks (below and above f_{tip}) were calculated.

$$S = \frac{L_1\text{-}L_2}{\log_2(f_2/f_1)}$$

Non-parametric analysis of variance was carried out by Kruskal-Wallis and post-hoc Mann-Whitney U testing using SPSS (IBM SPSS Statistics 23, NY, USA).

2.3 Results

All ears of barn owls (n = 14) showed SOAEs, with individual ears having between 9 and 16, on average 12.7 SOAEs. The pattern of SOAEs was unique to each ear. The comparison of right and left ears of each individual revealed no obvious correlation of the SOAE frequencies (f_{SOAE}). The f_{SOAE} ranged from 3.4 to 10.2 kHz. Figure 2.1 shows representative individual SOAE spectra. A total number of 178 SOAEs was observed. SOAE levels were clearly above the microphone noise (Fig. 2.2A). As an example, consider a small peak with a peak level at -20 dB SPL and a spectral width of 200 Hz. The peak level corresponds to 2 μPa. Thus, in the spectrum, the total area under the peak (L) is:

$$(\frac{\pi}{2}) \cdot 2^2 \cdot 200 = 1256\ \mu Pa$$

Hence the peak level (L) equals: $10 \cdot \log_{10}\left(\frac{1256}{20}\right) = 5\text{dB SPL}$, which is well above the noise floor for a bandwidth of 1 Hz (Fig. 2.2A). The noise level is thus substantially lower than the level of small peaks (Fig 2.1).

SOAEs overlapped at the base of the amplitudes and thus often formed a plateau that was well above the microphone noise floor and ranged in frequency from approximately 6.5 kHz - 10 kHz. Figure 2.2B shows that the emission peak width, determined from the Lorentzian curve fit, did not strongly correlate with SOAE level ($R^2 = 0.0034$). SOAEs were nearly regularly spaced on a linear frequency axis (Fig. 2.2C), with a median distance of 430 Hz (interquartile range of 179 Hz, range from 363 Hz to 542 Hz).

SOAE were stable within 1 dB over the time needed to obtain the recordings. Comparing f_{SOAE} before and after presentation of external tones (steps 1 and 3, see methods) showed maximal differences of around 300 Hz, and more typically less than 100 Hz.

2.3.1 Characteristics of suppression tuning curves

For 73 SOAEs, at least 2 dB of suppression was observed; most of these had a high f_{SOAE} and thus fell within the auditory fovea (>5 kHz). STCs were V-shaped and selectively tuned (Fig. 2.3A). The majority of the 73 SOAEs with STCs (71.2%) originated from the upper half of the auditory fovea, between 7.5 and 10 kHz. The tip of the STC could fall on either side of the emission frequency. In 76.7% of cases, the STC-tip was above the emission frequency.

The slope for each STC flank was measured between 3 and 23 dB SPL above the STC-tip. For 18 STCs, this suppression range was available on both flanks. The STC slope of the high-frequency flank (median: 179.9 dB/octave) was steeper than that of the low frequency flank (median: - 76.5 dB/octave). At higher levels, both the low- and high-frequency flank flattened out (Fig. 2.3A).

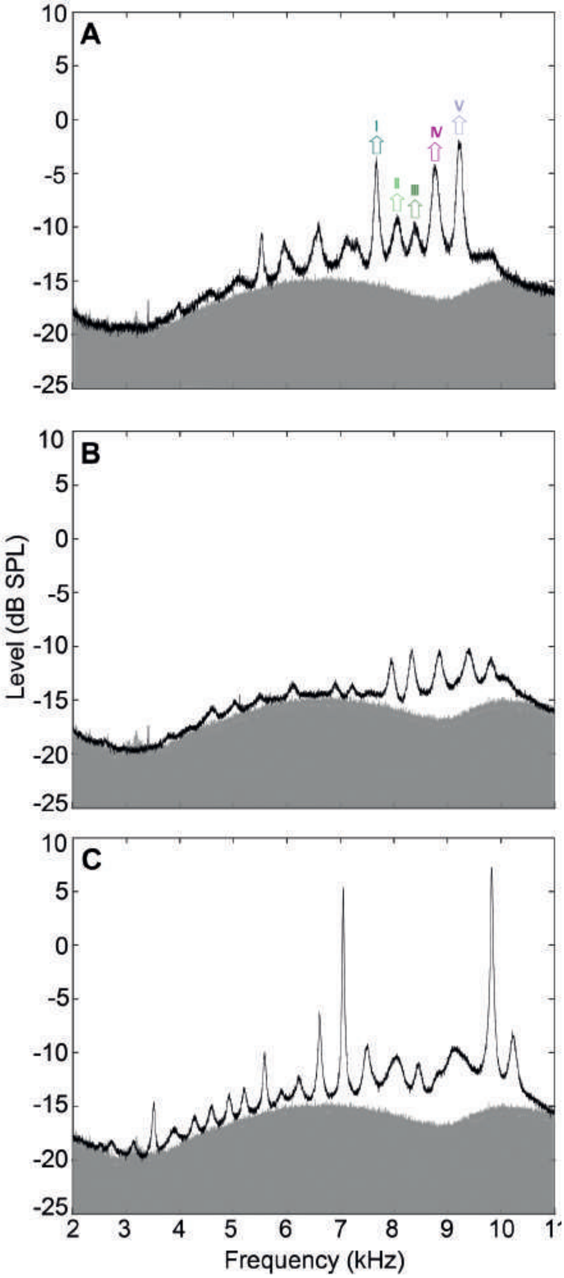


Figure 2.1. Three spectra of unsuppressed spontaneous otoacoustic emissions (SOAEs) of the barn owl. The spectral peaks correspond to the faint emission tones produced spontaneously from individual ears. Each ear showed a specific pattern of peak frequencies and amplitudes. In panel (A) 5 peaks are labeled: (I) at 7.67 kHz, (II) at 8.05 kHz, (III) at 8.40 kHz, (IV) at 8.77 kHz, and (V) at 9.23 kHz. The filled background shows the noise floor of the recording system. The spectral resolution is in 1-Hz bands.

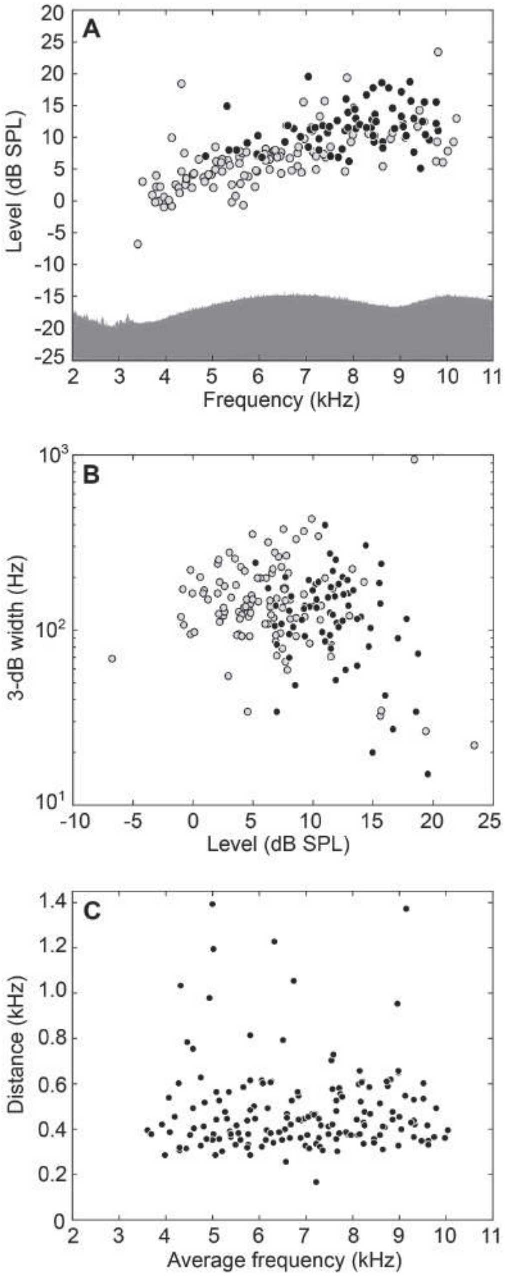


Figure 2.2. Characteristics of spontaneous otoacoustic emissions (SOAEs). Each circle corresponds to one SOAE peak (n = 178). For the SOAEs represented by black-filled circles in panels (A) and (B), suppression tuning curves were obtained (STCs, n = 73). (A) Relationship between SOAE frequency and SOAE level. The filled background shows the noise floor of the recording systemin 1-Hz bands. (B) SOAE peak width in relation to SOAE level. (C) Frequency distance between neighboring unsuppressed SOAE peaks (median distance = 430 Hz). The average frequency of each SOAE ($f_{average}$) was defined by the averaged spectrum of both unsuppressed recordings (see methods, step 1 and 3).

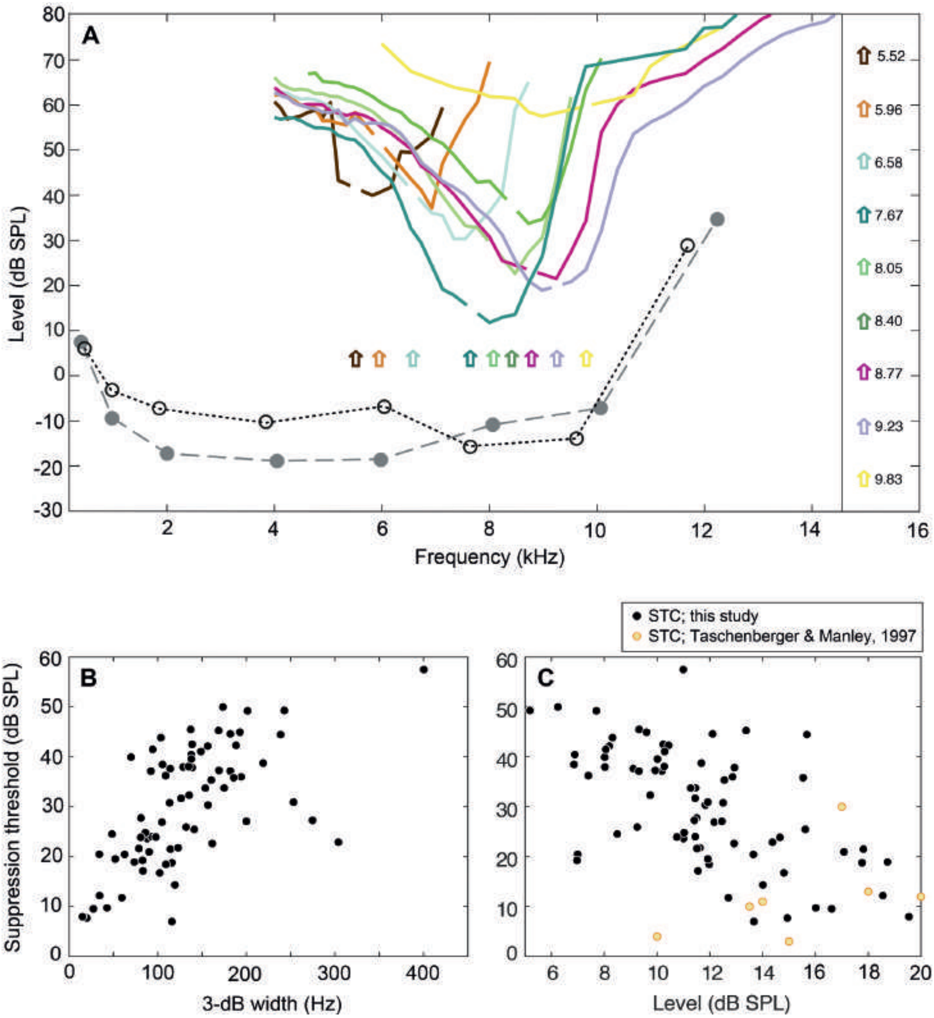


Figure 2.3. Suppression of spontaneous otoacoustic emissions (SOAEs). Suppression tuning curves (STCs) indicate the stimulus level needed for 2 dB suppression of the SOAE. (A) The STCs of one individual (spectrum in Fig. 2.1A). The arrows indicate the SOAE frequencies. The colors match the corresponding STC. The stimulus frequencies within 200 Hz of the unsuppressed spontaneous emission frequency were omitted (see main text) and appear as gaps in the STC. Behavioral thresholds in the barn owl are shown for reference, as black dotted lines (Krumm et al., 2017) and grey dashed lines (Konishi, 1973). (B) STC threshold as a function of unsuppressed SOAE width. (C) STC threshold as a function of unsuppressed emission level. Black-filled circles indicate STCs from this study (n = 73) and filled orange circles data from Taschenberger and Manley (1997; n = 7). Note that 10 dB were added to the SOAE levels from Taschenberger and Manley (1997), to correct for the different methods in level estimation between both studies.

Tuning curve threshold

The thresholds of the 2 dB STCs varied from 7.0 to 57.5 dB SPL, with no trend across $f_{average}$ ($R^2 = 0.05$; $p = 0.07$). Figure 2.3B shows that narrower SOAEs were suppressed by external tones of lower sound pressure levels than spectrally broader SOAEs ($R^2 = 0.39$; $p < 0.001$). Furthermore, SOAEs with relatively lower levels required a higher sound level for suppression, whereas larger SOAE levels were suppressed by tones of lower sound pressure levels (Fig. 2.3C; $R^2 = 0.36$, $p < 0.001$). A comparison of the methods to derive SOAE levels of Taschenberger and Manley (1997; peak level) and our study (area under the peak) was carried out on our new data, to assess the difference that it potentially makes to the results. Peak levels were typically 10 dB lower. In order to show SOAE levels of both studies in a comparable way, we therefore added 10 dB to all the Taschenberger and Manely (1997) data (Fig. 2.3C).

In order to compare the STCs to neural tuning curves (TCs) in the same species, data from two previous reports were plotted together with the results of the present study (Fig. 2.4A). Taschenberger and Manley reported a median STC threshold of 11 dB SPL ($n = 8$), and the median neural TC threshold was 14 dB SPL ($n = 246$; Köppl, all data shown in Fig. 2.4A). In the present study, a higher median STC threshold was obtained (30.80 dB SPL, $n = 73$). A Mann-Whitney U test revealed significant differences ($p < 0.005$) between the STC thresholds of this study compared to suppression thresholds reported in 1997 by Taschenberger and Manley ($U = 60$) and this current study compared to neural TC thresholds reported by Köppl (1997a), b, and unpublished results ($U = 2633.5$).

Tuning curve Q_{10dB}

The STC median Q_{10dB} value was 4.87 ($n = 73$). Q_{10dB} was independent both of SOAE level ($R^2 = 0.0012$; $p = 0.77$) and of SOAE width ($R^2 = 0.012$; $p = 0.36$). Q_{10dB} values of this study were compared to previous suppression- and neural TCs (Fig. 2.4B). Taschenberger and Manley reported a median Q_{10dB} of 8.2 ($n = 8$) and the neural TC dataset of Köppl (1997a, b and unpublished results) revealed a median Q_{10dB} of 5.7 ($n = 218$). The Mann-Whitney U test showed significant differences between the Q_{10dB} values of the current STC study and the STCs study published 1997 by Taschenberger and Manley ($U = 23$, $p < 0.05$) and between the current study and the neural TCs published by Köppl (1997a, b), and unpublished results ($U = 10935$, $p < 0.05$).

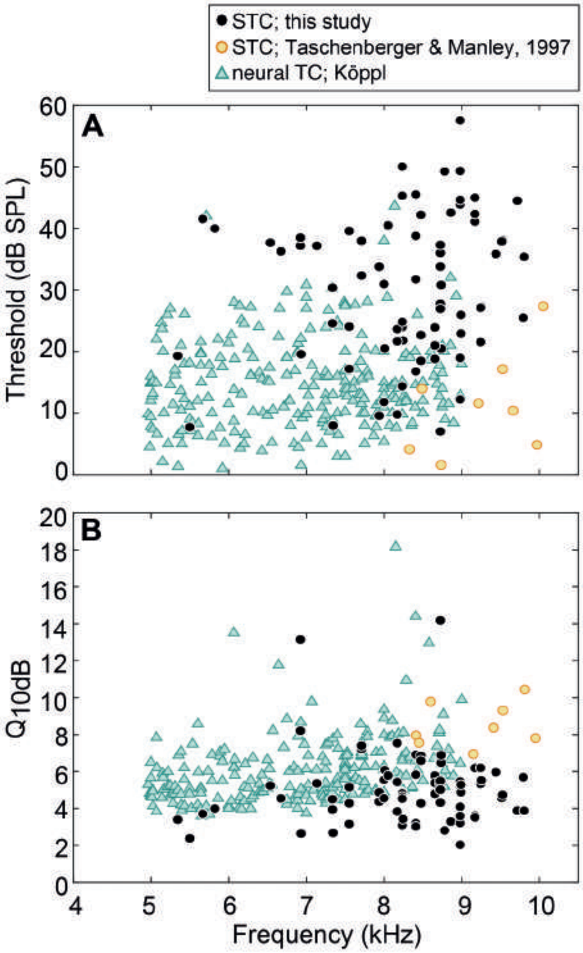


Figure 2.4. Comparison between tuning curves of SOAE suppression and auditory-nerve single-unit recordings. (A) Thresholds of STCs and neural TCs as a function of tuning curve tip frequency. (B) The filter quality factor Q_{10dB} of STCs and neural TCs as a function of tuning curve tip frequency. SOAE suppression tuning curves: filled black circles (this work) and filled orange circles (Taschenberger and Manley, 1997). Neural tuning curves: filled turquoise triangles (Köppl, 1997a, b, and unpublished results), for the frequency range from 5 to 10 kHz.

2.4 Discussion

2.4.1 General characteristics of the SOAEs

As in mammals, SOAEs are rare in birds. The barn owl is thus far the only known bird species showing SOAEs. Considering that SOAEs have been reported in all groups of land vertebrates, it is assumed that these emissions are caused by a symplesiomorphic active process that evolved in ancestral species and constitutes a fundamental feature of all inner ears (Manley, 2001). In mammals, but not in birds, it is further assumed that emission energy originates by the action of prestin (Dallos et al., 2008; Xia et al., 2016). The emission patterns are specific for each species and individual (Manley, 2001), suggesting that the species' and individual's morphology affects spectral patterning. In birds, the degree of interaural coupling in general decreases with both increasing head size and increasing frequency. For the barn owl, it has been shown that interaural attenuation increases to values of minimally 35 dB at 7 kHz and above (Moiseff and Konishi, 1981; Palanca-Castan et al., 2016). Thus, most or all of the measured SOAEs are not expected to interact between the ears and we also found no evidence for such interactions. Many studies have shown that the widespread phenomenon of SOAE suppression relates to individual frequency tuning properties (Manley and van Dijk, 2008).

In this study, many more SOAEs per ear, in particular ones with smaller levels, were recorded than 20 years ago by Taschenberger and Manley (1997; comparison in Fig. 2.3C). This is presumably due to the higher sensitivity of the equipment used.

If SOAE in any individual ear did shift in frequency, all SOAE shifted in the same direction, suggesting a common influence such as minor variations in body temperature (that have large effects, see Taschenberger and Manley, 1997) or possibly changes in tonic efferent activity (Manley et al., 1999).

The distance between neighboring SOAEs was near 430 Hz in all frequency ranges and across ears (Fig. 2.2C and Supplementary Fig. 2.1). This contrasts with emission spectra in humans, where the spacing between SOAE peaks increases with increasing frequency of the neighboring peaks (reviewed in Shera, 2003). The spacing in human SOAE spectra presumably reflects standing-wave conditions for which backward and forward traveling waves in the cochlea can combine to produce a standing wave on the basilar membrane. In lizard SOAE spectra also, the spacing generally increases with the peak frequency (Manley et al., 2015). In birds, including barn owls, sharply tuned traveling- or standing waves presumably do not exist on the basilar membrane, since in pigeons and chickens only broadly-tuned traveling waves without evidence for nonlinear amplification were observed (Gummer et al., 1987; Xia et al., 2016). This is in apparent contrast with independent evidence for cochlear amplification and nonlinear behavior, such as the high sensitivity and sharp tuning of auditory nerve fibers, otoacoustic emissions, and active motile processes in hair cells (e.g.: Manley, 2001; Peng and Ricci, 2011; Beurg et al., 2013). Although membrane channel densities and kinetics (electrical tuning) contribute to sharp frequency tuning, this component fades towards the upper frequency range of bird hearing, above several kHz (Wu et al., 1995), i.e. in the frequency range of particular interest in the barn owl.

2.4.2 SOAE suppression by external tones

In all classes of terrestrial vertebrates so far studied, SOAEs have been shown to be sensitive to the presence of external tones, especially near their peak frequency. In barn owls, also, SOAE level was suppressed by external tones, depending on the frequency distance between the external stimuli and the SOAE and on stimulus level. Stimuli closer in frequency to the SOAE had a larger suppressive effect than those further away, and tones of higher level were more suppressive than those of low level. Thus, the typical V-shaped STCs were observed. The suppression tuning curves obtained here were similar in their shape to those observed in the earlier study of barn owls (Taschenberger and Manley, 1997).

Tuning curve tip and frequency pushing and pulling

In humans, the tip frequency of STC is consistently found above the SOAE frequency (Schloth and Zwicker, 1983; Zizz and Glattke, 1988; Manley and van Dijk, 2016: 4.5% higher). In our study, the most effective suppressor stimulus in owls was either below or above the SOAE peak frequency, with a tendency that STC-tips lay above emission frequency. Due to the analysis procedure, it was not possible to fully evaluate the tip region of the STCs, i.e. stimulus frequencies within ±100 Hz of the emission frequency.

Geisler et al. (1990) described a mammalian cochlear model that examined the source of SOAEs and the shift in STC-tip frequency towards higher frequencies. This model might not be applicable to all vertebrates with SOAEs (e.g.: lizards and barn owls), as it requires mammal-like traveling waves and a mammalian active mechanism; consequently, other models have to be considered (e.g.: Bergevin and Shera, 2010). Earlier SOAE suppression studies in other species, such as lizards (Köppl and Manley, 1994; Manley et al., 1996; Manley, 2004, 2006), described f_{SOAE} changes caused by external tones. Generally, the f_{SOAE} shifted away from the stimulus frequency (frequency “pushing”), especially when the stimulus frequency was above the emission frequency. Stimuli of greater sound pressure level and frequency nearer the emission frequency increased the f_{SOAE} shift up to several hundred Hz (Köppl and Manley, 1994; Manley et al., 1996; Manley, 2004). Human SOAEs can also be both pushed away from or pulled towards an external stimulus (Long, 1998; Baiduc et al., 2014; Manley and van Dijk, 2016). This SOAE shift is, however, very much smaller in humans than in lizards. Presumably, human SOAEs are frequency stabilized by the standing-wave mechanisms discussed above.

Interestingly, in barn owls we did not observe consistent pushing or pulling of the SOAEs that depended on stimulus level or frequency (Supp. Fig. 2.2). It is currently not clear why barn owl SOAEs are relatively stable in frequency when being suppressed by external tones, despite the presumed absence of standing waves that may serve as a stabilizing mechanism.

Tuning curve slopes and secondary side-lobes

The asymmetric shape of STCs, with steeper slopes for the high frequency flank (human: Zizz and Glattke, 1988, Manley and van Dijk, 2016; macaque monkey: Martin et al., 1988;

most lizards: Manley and van Dijk, 2008) or the lower frequency flank (some lizards: Manley, 2006) describes an almost universal phenomenon of asymmetrical inner-ear tuning. Comparable tuning curves for neural tuning (lizards: Manley et al., 1990; Köppl, 1997a; Manley, 2001) and STCs within the same species (e.g.: Martin et al., 1988; Manley and van Dijk, 2016) have been reported. Consistent with neural tuning curves of the barn owl (Köppl, 1997a), SOAE STCs were characterized by a steeper slope of the higher frequency flank.

Unlike other species, such as humans (Manley and van Dijk, 2016), macaque monkey (Martin et al., 1988), and many lizards (Köppl and Manley, 1994; Manley, 2001), the STCs of barn owls lacked very sharp secondary sensitivity tips on the high-frequency flank of STCs (Taschenberger and Manley, 1997) and of neural TCs (e.g.: Köppl, 1997a). Consistent with Taschenberger and Manley (1997) we found, however, that the high-frequency flank of some STCs flattened out towards the high suppressor levels, something which was never observed in neural TCs. In humans, the side-lobes were attributed to the interactions between the suppressing stimulus and the SOAE standing wave (Manley and van Dijk, 2016).

The absence of secondary suppression lobes in the barn owl can be interpreted as standing waves not being present. This may reflect expected differences in the cochlear mechanics of the barn owl compared to mammals. Note that these secondary minima were also seen in neural tuning curves in the bobtail and other lizard species (e.g.: Manley, et al., 1988). However, the side-lobes of STCs and neural tuning curves in lizards cannot be caused by standing waves, as suggested for humans, as there are no traveling waves on the basilar membrane (e.g.: Manley, et al., 1988). The inconsistent presence of side-lobes in suppression tuning curves and neural tuning curves suggests different inner ear tuning mechanisms in mammals, birds and lizards.

Behaviorally obtained hearing thresholds of the barn owl indicated sensitive hearing between 200 Hz and 12 kHz (Konishi, 1973; Krumm et al., 2017). However, SOAEs were also suppressed by higher-level (>~55 dB SPL), high-frequency external sounds above the behaviorally tested hearing range. High-frequency STC flanks reached up to the very highest frequency of the owl’s hearing range and even extended it (Fig. 2.3A). Consequently, we suggest that behavioral hearing threshold estimation should include frequencies above 12 kHz.

Tuning curve tip thresholds and their relation to SOAE width and level

An unexpected observation was that both SOAE level and width were related to STC-tip threshold, such that narrower and larger SOAE suppressed more easily, with lower thresholds (Fig. 2.3B and C). At present, we can only speculate on the origin of these correlation by considering simple oscillator models (Stratonovich, 1967). The models tend to suggest a relation between oscillator amplitude and suppression threshold that is reverse to what has been observed here: in the oscillator model the effectiveness of an external force (amplitude E) to modulate a self-sustained oscillation (amplitude A) always depends on the ratio E/A. The larger the oscillator amplitude A, the stronger the external force E is needed to affect the oscillator's behavior. In the current work, the reverse appears to be true. The relation between suppression threshold and the ratio E/A of an external suppressor tone (E) and the oscillation amplitude (A), assumes that the internal noise level, to which the oscillator is exposed, is relatively constant. Specifically, the noise level is considered to be constant across SOAEs with various oscillation amplitudes. This assumption appears to be approximately correct for human SOAEs, where a negative correlation between SOAE width and level was found (Talmadge et al., 1993; van Dijk et al., 2011). However, in the barn owl, SOAE peak height and width are not significantly correlated (Fig. 2.2B). As a consequence, the internal noise of the SOAE oscillator is not at a constant level across SOAE peaks. The oscillators internal noise counteracts its synchronization to an external tone. Thus, less internal noise implies easier synchronization with lower suppression thresholds. Consistent with this view, relatively narrow SOAEs have low suppression thresholds (Fig. 2.3B).

The STC results of the present study were plotted together with the already published STCs and neural TCs of the barn owl (Fig. 2.4A). Between 5 and 10 kHz, both STC measurements (Taschenberger and Manley, 1997) and TCs of single auditory nerve fibers (Köppl, 1997a, b, and unpublished results) show similar best thresholds. In the present study, a higher STC threshold was obtained which, however, falls within the range of the previously observed thresholds (STC: 1.55-27.33 dB SPL, neural recordings: 1-43.6 dB SPL). This is plausibly explained by the negative correlation between SOAE suppression threshold and SOAE level: weak SOAEs have high suppression thresholds (Fig. 2.3C) and the more sensitive recording equipment allowed the recording of many more small SOAEs. Consequently, overall SOAE suppression thresholds are higher in the current study when compared to Taschenberger and Manley (1997).

STC sharpness: Q_{10dB}

Here, the current data are compared to previous reports of STCs (Taschenberger and Manley, 1997) and neuronal TCs Köppl (1997a, b and unpublished results) of the barn owl (Fig. 2.4B), within the overlapping frequency range from 5 to 10 kHz. The Q_{10dB} values were similar, but lower in the current study.

Another difference to previous findings was the absence of any frequency dependence on tuning sharpness in our data. Köppl (1997a) showed that barn owl eighth-nerve axons were narrowly tuned, even at SPLs much above CF threshold. The mean neural Q_{10dB} increased with CF according to a power law from 1.7 at 0.5 kHz to 7.25 at 9 kHz (Köppl, 1997a). Similarly, in behavioral data, the auditory filter bandwidth increases within the auditory fovea (Dyson et al., 1998). In contrast, the SOAE suppression measurements described here did not reveal such a trend; a regression across SOAE-STC sharpness data was flat (Fig. 2.4B).

In humans and in lizards, there is a clear trend for STC tuning sharpness to increase with frequency (Manley et al., 2015). If this reflects the logarithmic distribution of frequencies in the tonotopy of the papillae of these species, then the lack of such an increase in the barn-owl data simply reflects the almost linear distribution of approximately 80% of the frequency range of its cochlea (Köppl et al., 1993).

In summary, STCs are similar to neural TCs in some details but were, on average, less sensitive and less sharply frequency tuned (Fig. 2.4B), especially at high sound levels. For several species, Q_{10dB} values of SOAE-STCs were found to be equivalent to neural tuning curves derived from auditory nerve fiber recordings (e.g.: compare barn owl: Taschenberger and Manley, 1997 with Köppl, 1997a; macaque: Martin et al., 1988 with Shera et al., 2011, lizards: Manley et al., 1990 with Köppl and Manley, 1994). However, the current study does not confirm this impression of detailed similarity between neural and suppression TCs, despite apparent support from the smaller sample in the work of Taschenberger and Manley (1997). This cannot be explained by sampling biases for different types of TCs. In birds, including the barn owl, there is no evidence for populations of auditory nerve fibers with distinct physiological properties. In particular, there are no subgroups distinguished by spontaneous discharge rate, since spontaneous rates show a monomodal distribution. There is also no correlation between spontaneous rate and other physiological properties such as response threshold or tuning sharpness (e.g.: Köppl, 1997a, 2011).

In mammals under ideal recording conditions (Sellick et al., 1982; Rhode, 1995; Narayan et al., 1998), tuning at the basilar membrane level matches recordings of single auditory nerve fibers. This is unlikely to be the case in birds. Although equivalent measurements are not available for barn owls, in both chicken and pigeon, basilar-membrane motion showed poorer frequency tuning than auditory-nerve fibers, and no clear evidence for active amplification (Gummer et al., 1987; Xia et al., 2016).

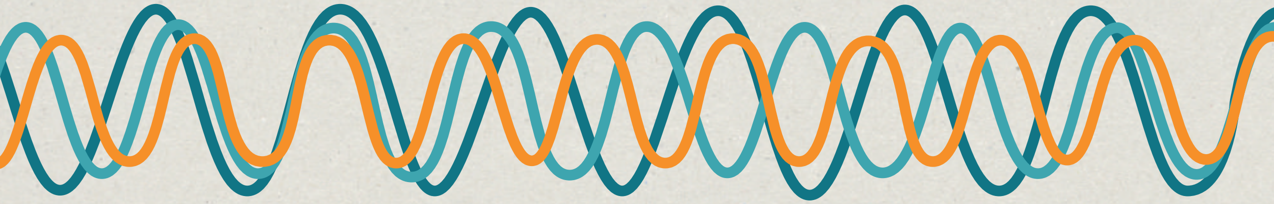
2.5 Conclusions

In this study, SOAEs of both ears in 7 barn owls were recorded and suppressed by pure-tone stimulation. The frequency separation between neighboring peaks was approximately constant across frequency. Unlike in humans and lizards, secondary dips of suppression on the high-frequency flanks of STCs were not found. This suggests that peripheral processing of SOAE suppression in birds - or at least in the barn owl - differs in this respect from that of lizards and humans. The negative correlation between SOAE width and sensitivity to suppression and the constant frequency spacing to SOAE peaks are likely to be indicators of fundamental properties of the owl's inner ear.

Acknowledgements

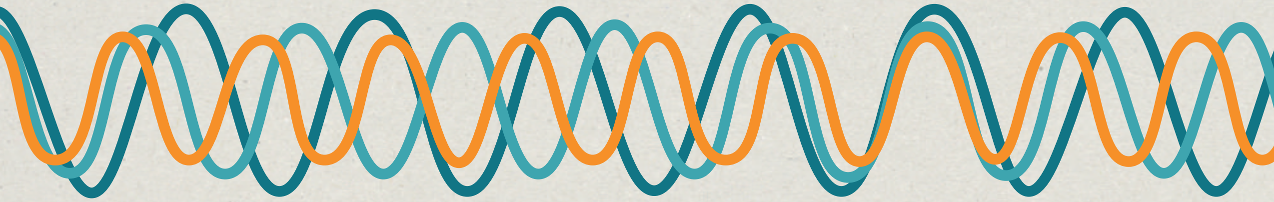
We thank Paolo Toffanin for programming. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie-COFUND (EGRET, Grant No. 661883) and the DFG Cluster of Excellence EXC 1077/1 "Hearing4all".

Supplementary data to this article can be found online at: https://doi.org/10.1016/j.heares.2019.107835.



Chapter 3

Frequency selectivity of tonal language native speakers probed by suppression tuning curves of spontaneous otoacoustic emissions



S. Engler, E. de Kleine, P. Avan, P. van Dijk

This chapter has been published as Engler et al., 2020. Hearing Research 398, 1-8.

Abstract

Native acquisition of a tonal language (TL) is related to enhanced abilities of pitch perception and production, compared to non-tonal language (NTL) native speakers. Moreover, differences in brain responses to both linguistically relevant and non-relevant pitch changes have been described in TL native speakers. It is so far unclear to which extent differences are present at the peripheral processing level of the cochlea. To determine possible differences in cochlear frequency selectivity between Asian TL speakers and Caucasian NTL speakers, suppression tuning curves (STCs) of spontaneous otoacoustic emissions (SOAEs) were examined in both groups. By presenting pure tones, SOAE levels were suppressed and STCs were derived. SOAEs with center frequencies higher than 4.5 kHz were recorded only in female TL native speakers, which correlated with better high-frequency tone detection thresholds. The suppression thresholds at the tip of the STC and filter quality coefficient Q_{10dB} did not differ significantly between both language groups. Thus, the characteristics of the STCs of SOAEs do not support the presence of differences in peripheral auditory processing between TL and NTL native speakers.

3.1 Introduction

Languages can be differentiated into tonal (TL) or non-tonal (NTL). Several studies addressed a link between native language and the acuity of pitch perception. In TL, such as Chinese, pitch changes signal different lexical meanings of the same word. Therefore, in TL, the precise perception of pitch alterations is essential for the understanding of lexical content. It is not surprising that native speakers of TL pay more attention to pitch changes (Braun and Johnson, 2011) and outperform native NTL speakers in pitch interval discrimination (Pfordresher and Brown, 2009; Hove et al., 2010; Giuliano et al, 2011). Producing and perceiving TL-cues may enhance pitch perception (Pfordresher and Brown, 2009; Giuliano et al., 2011) and production (Deutsch et al., 2004). These findings indicate that the individual linguistic background potentially affects pitch perception to some degree.

Depending on native language background, different brain areas become active during a discrimination task of linguistic stimuli (Gandour et al., 1998; 2000). In general, language processing is lateralized to the left-brain hemisphere, whereas tonal pitch processing takes place in the right hemisphere (Zatorre et al., 1994; 2002). Only in TL native speakers, areas of the left brain-hemisphere are activated during pitch processing in a linguistic context (Gandour, 1998). This hemispheric asymmetry might be even expected, as complex linguistic cues are predominantly processed in this hemisphere. When discriminating lexical tones in linguistic contexts, TL native speakers show activation in Broca's area, whereas NTL native speakers do not (Gandour et al., 1998, 2000; Wong et al., 2004). Thus, the activation of particular brain areas, depending on the cues of a language and the listener's language experience, indicate specific differences when processing speech.

Language experience may, however, also influence fundamental auditory processing of sounds with no linguistic content (Salmelin et al., 1999; Vihla et al., 2002). For example, an increase in absolute pitch prevalence can be reached by some kind of training effect due to TL acquisition (Deutsch et al., 2004; 2006; Pfordresher and Brown, 2009). In fact, native speakers of TL also have enhanced pitch perception also in non-linguistic contexts, and outperform NTL control groups (Deutsch et al., 2006; Krishnan et al., 2009; Giuliano et al., 2011). Deutsch et al. (2006) described that absolute pitch perception of musically trained TL native speakers is even more precise than that of musically trained subjects with a NTL background. It is unclear, however, whether these enhanced perceptual abilities of TL native speakers are directly related to cochlear frequency selectivity. Fundamental inner ear properties can potentially cause differences in frequency selectivity and pitch perception.

Otoacoustic emissions (initially described by Kemp, 1978) allow the non-invasive and objective measurement of frequency selectivity. In the absence of any external stimulation, sounds can be emitted by the ear. These sounds are termed spontaneous otoacoustic emissions (SOAEs). SOAEs are continuous sinusoids with small fluctuations in frequency and level. They can be recorded by placing a sensitive miniature microphone in the ear canal. In humans, otoacoustic emissions are believed to be produced by outer hair cell activity and thus, may indicate healthy hair cell properties (Brownell et al., 1985). The presence of SOAEs is not rare, as approximately 70% of young and normal-hearing humans emit them (Talmadge et al., 1993). SOAEs are, however, not essential for an adequate acoustic perception. Interestingly, human SOAE occurrence differs between genders, with females having a higher SOAE prevalence than males. Moreover, SOAE prevalence differs between ethnic groups, with Asians expressing more SOAEs per ear than Caucasians (Whitehead et al., 1993). What causes these differences in SOAE occurrence remains so far speculative, but could indicate differences in ear properties between Asians and Caucasians.

External tone stimuli have characteristic and frequency selective suppression effects on SOAEs. Suppression tuning curves (STCs) can be derived by measuring the suppression of a single emission peak for various frequencies and levels of the external tone. STCs of SOAEs allow objective and non-invasive estimation of the cochlear frequency selectivity (Schloth and Zwicker, 1983).

The general approach of the current study can be compared to the research of McKetton et al., 2018, who investigated the prevalence of SOAEs and cochlear tuning in participants with and without absolute pitch perception. We examined the cochlear frequency selectivity of Asian subjects with a TL mother tongue and Caucasian subjects with native NTL background, using STCs of SOAEs. We evaluated whether human frequency selectivity at the cochlear level differs systematically between ethnic groups with different native language experience.

3.2 Material and methods

3.2.1 Participants

The recordings were carried out in healthy adults, aged between 18 and 31 years. All participants were screened for SOAE occurrence and had normal hearing thresholds at the emission frequency with pure tone thresholds of ≤25 dB hearing level (HL). Participants self-reported a clear Asian or Caucasian heritage with either TL or NTL native language experience (respectively). Eight out of 17 participants with TL background and three out of 16 with NTL background were musically trained. None of them was a professional musician (definition: Micheyl et al., 2006).

This study was approved by the Medical Ethics Committee of the University Medical Center Groningen, Netherlands (Letter of March 11th 2014, METc 2014.099). The Comite d’Evaluation Ethique de l’Inserm (Letter of March 21st 2019, CD/EB 19-034) approved this study in France. The study was conducted in accordance with the Declaration of

Helsinki and applicable laws. All participants gave written, informed consent, and received a modest financial compensation for their participation.

3.2.2 Recording procedure

In each participant both ears were examined for the presence of SOAEs. The hearing thresholds of both ears were measured by using pure tone audiometry (Audiosmart, Echodia, Clermont Ferrand, France) at: 0.25, 0.5, 1, 2, 4, and 8 kHz. In each ear in which SOAEs were present, the recording procedure encompassed three main steps: 1) A two-minute SOAE recording without external stimuli. 2) A one-hour suppression measurement, presenting tones over a large number of levels and frequencies, in quasi-random order (exact test duration depended on the number of stimuli presented). 3) A two-minute SOAE recording, equivalent to step 1.

The measurements were conducted at two locations: The University Medical Center Groningen (UMCG, Groningen, Netherlands) and the University Clermont Auvergne (UCA, Clermont Ferrand, France). At the UMCG the measurements were carried out in a double-walled, sound-attenuating chamber (Industrial Acoustics Company, Niederkrüchten, Germany). At the UCA, the measurements were carried out in a quiet office.

SOAE recording

An occluding soft foam ear plug, including the Etymotic ER-10C microphone-speaker system (Etymotic Research, Inc., Elk Grove Village, IL, USA), was placed in the external ear canal. The microphone output was amplified by 40 dB, using the Etymotic ER-10C preamplifier. During the measurements at the UMCG (except for one ear), an additional amplification of 20 dB was applied by using the Stanford Research System (Stanford Research Systems Inc, model SR 640, CA, USA). SOAEs were monitored by feeding the amplified signal into a spectrum analyzer (SRS Inc., model SR 760).

An audiofire AD/DA converter was used to record the microphone signal on the computer disc and to generate the tone stimuli. At the UMCG (Netherlands) we used the Motu 624 (MOTU Inc., MA, USA) for AD/DA conversion, while at the UCA (France) the ESI U 24 XL (ESI Audiotechnik GmbH, Leonberg, Germany) was used. Stimulus generation and response recording was controlled by custom routines developed with Matlab software (MathWorks Inc., 2016a, Natick, MA, USA).

Emission peaks in the time-averaged spectrum that exceeded the noise floor and that were suppressible by external tones were identified as SOAEs. These SOAEs were individual for each ear. The emission recording with the best signal-to-noise ratio was used to calculate the SOAE frequency (f_{SOAE}), the emission width and level. We excluded small frequency components that were not amenable to the Lorentzian curve fit from further analysis. SOAE suppression by at least 3 dB by external tones lower than 70 dB was required as a further inclusion criterion for the STC analysis.

Stimulus presentation

Stimulus tones of different frequencies and levels were presented in a quasi-random order to investigate the suppressive effect of external tones on the SOAEs. The stimulus frequencies were chosen to cover the range in which SOAEs were detected. In most cases, the suppression frequency varied from 500 Hz to 10 kHz in 1/16 octave steps. The stimulus levels ranged from 0 to 70 dB SPL in 3 dB steps. Thus, the total number of stimuli was 1587. Each stimulus had a duration of 1.2 s (with a 10-ms cosine rise and fall time). For each stimulus tone, a segment of 1.5 s of the microphone signal was recorded and stored. The SOAE recording started 150 ms prior to the tone onset and ended 150 ms after the tone offset. The center one second of this recording was evaluated for suppression effects of the external tone on the SOAE.

3.2.3 Data analysis

To determine the effect of a single external tone on an SOAE, the entire 1-second interval (described above) was analyzed, during which the stimulus tone was present. To include the SOAE, but exclude the stimulus tone and its harmonics, a tonal signal with a frequency equal to the stimulus plus two higher harmonics was fitted to the time-domain of the recorded signal. The resulting fit was subtracted from the recorded signal, creating a residual. A zero-phase band-pass filter with a level response determined by the emission frequency (f_{SOAE}) and the width of the filter (Δf) was applied to the residual to isolate the SOAE of interest:

$$A(f) = \left[1 + \frac{(2[f - f_{SOAE}])^{8}}{\Delta f} \right]^{-\frac{1}{2}}$$

The center frequency of the 60 Hz wide filter was placed at the unsuppressed f_{SOAE}. Subsequently, the Hilbert phase of the filtered signal was used to compute the average emission frequency during the 1-s recording segment. The filter procedure was repeated, with a filter center frequency that now equaled this computed SOAE frequency, and a width narrowed to 10 Hz, for further noise reduction. For stimulus tones closer than this 10 Hz window, SOAE suppression was not assessed. The SOAE level was determined as the average Hilbert envelope during the 1-s interval. This procedure was repeated for each f_{SOAE} and the characteristics of the stimulus, thus creating a full frequency matrix of SOAE levels. Each element of the matrix contained the individual SOAE level for a given stimulus level and -frequency. Sound fragments that contained artefacts (resulting from movements, swallowing, etc.), as determined by an artifact level crossing paradigm, were ignored.

In the further analysis, we only included SOAEs if they were sufficiently strong relative to the noise floor, to create a tuning curve for 3 dB suppression. STCs were characterized by all relevant suppressor-tone frequencies and levels at which the emission reached 3 dB attenuation. Moreover, STCs that consisted of less than 4 data points or were very noisy were excluded from the analysis. The weakest stimulus that produced 3 dB suppression was referred to as the suppression threshold, with a corresponding tip-frequency (f_{tip}) of the suppression tuning curve. The STC sharpness was calculated by the filter quality factor Q_{10dB}. This factor is defined as the ratio between f_{tip} and the width of the tuning curve, at 10 dB above the tip (Δf_{10dB}):

$$Q_{10dB} = \frac{f_{tip}}{\Delta f_{10dB}}$$

Tuning curves slopes were evaluated for the lower and the higher frequency flank according to the threshold level at f_{tip} (L_1) and 10 dB above threshold (L_2). The corresponding frequencies (f_1 and f_2) were interpolated. The slope S is defined as:

$$S = \frac{L_1 - L_2}{\log_2(f_2/f_1)}$$

3.3 Results

We included 17 TL native speakers (Chinese) of Asian heritage of whom 13 were female and four were male (see table 3.1). The group of NTL native speakers (Dutch, German) consisted of 16 subjects of Caucasian heritage; 12 of them were female and four were male. The median age of TL native speakers was 22.2 years, NTL native speakers had a median age of 21.2 years.

All SOAEs included in this analysis (n = 181) showed 3 dB STCs, necessary to evaluate the frequency selectivity. SOAE levels were clearly above the microphone noise. These SOAEs were stable over the time needed to obtain the suppression measurement. In both tested language groups, the number of SOAEs varied between individuals and ears. We recorded 95 SOAEs in 27 ears of Asian TL native speakers and 86 SOAEs in 23 ears of Caucasian NTL native speakers. The majority of the participants (n = 28) were tested in the Netherlands. Five female TL native speakers were tested in France.

Table 3.1. Overview of evaluated spontaneous otoacoustic emissions (SOAEs) with 3 dB suppression tuning curves (STCs) for both language groups. Number of participants, number of ears producing SOAEs, and number of SOAEs are indicated per group.

	Native language group	
Evaluated STCs	**Tonal**	**Non-tonal**
Total $N_{Participant}$	17	16
Total N_{Ear}	27	23
Total N_{SOAE}	95	86
Female participants		
$N_{Participants}$	13	12
N_{Ear}	22	18
N_{SOAE}	88	80
Male participants		
$N_{Participants}$	4	4
N_{Ear}	5	5
N_{SOAE}	7	6

3.3.1 Spontaneous otoacoustic emissions (SOAEs)

In both native language groups, SOAEs were more often recorded in females than in males (see Table 3.1). In the TL group, we recorded SOAEs in 22 ears of females and in five ears of males. In the NTL group, we recorded SOAEs in 18 ears of females and five ears of male participants. Moreover, females did not only tend to have SOAEs more commonly, they also had more emissions per emitting ear. In native speakers of TL, SOAEs of females represent 93% of all recorded SOAEs (n = 95). In the NTL group, also 93% of all recorded SOAEs (n = 86) were recorded in females.

The frequency distribution of the SOAEs is shown in Figure 3.1, and was similar between both language groups (Kolmogrorov-Smirnov test, p = 0.176). TL native speakers had SOAEs between 0.63 and 8.53 kHz (median: 1.84 kHz). In the NTL native speakers, SOAEs ranged from 0.60 kHz to 4.47 kHz (median: 1.84 kHz). Thus, the SOAEs of TL native speakers were recorded in a wider frequency range towards the higher frequencies. In seven ears (26%) of TL native speakers, SOAEs with frequencies larger than 4.5 kHz were recorded. The hearing sensitivity of these ears was not exceptionally good at these frequencies. However, in general, TL native speakers had relatively good hearing thresholds over the frequency range from 2 to 8 kHz with mean audiometric thresholds between 0.4 and 2.2 dB HL (Fig. 3.1A). An independent sample t-test with Bonferroni correction revealed significant hearing threshold differences between both language groups at 0.5 kHz (p < 0.001), 1 kHz (p = 0.002), and 8 kHz (p < 0.001).

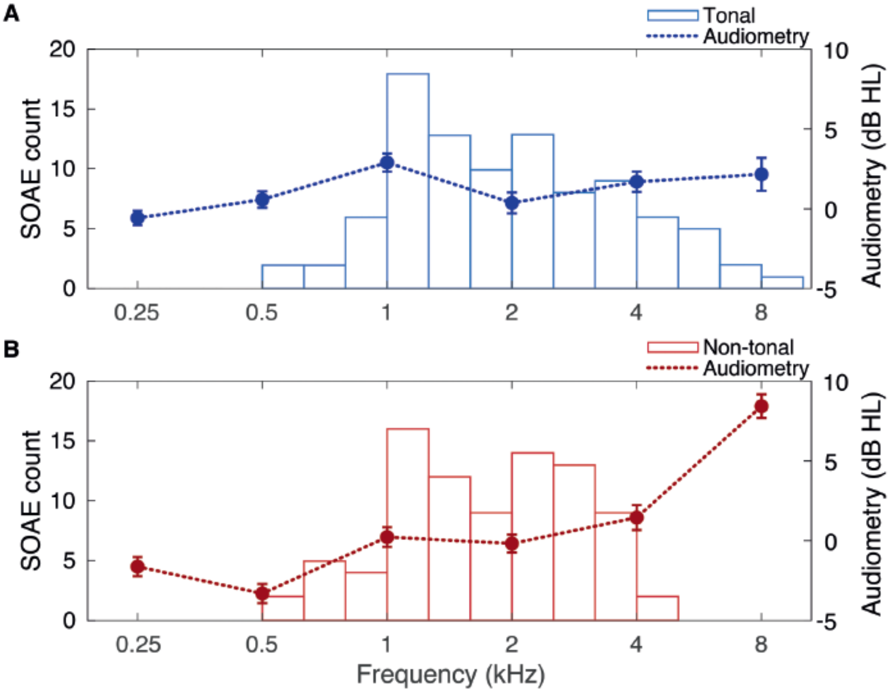


Figure 3.1. Distribution of SOAEs and hearing thresholds. Panels show the number of SOAEs per frequency band and the corresponding average pure tone thresholds, per group. Each SOAE count corresponds to one emission peak from which STCs were obtained (n = 181). (A) SOAEs of TL native speakers covered a frequency range from 0.63 to 8.52 kHz (n = 95). Panel (B) shows the SOAEs of NTL native speakers (n = 86). SOAEs of this language group cover a range from 0.60 to 4.46 kHz. Mean pure tone thresholds are plotted with standard error bars. Significant differences in hearing threshold between both language groups were present at 0.5, 1, and 8 kHz.

Both language groups had similar SOAE levels (Kolmogrorov-Smirnov test, $p = 0.251$), as can be seen in Figure 3.2A. Both language groups show differences in emission levels across frequencies that were related to the higher f_{SOAEs} recorded in the TL group. In TL native speakers, no correlation between the SOAE level and f_{SOAE} was found ($R^2 = 0.02$; $p = 0.18$), also when evaluating f_{SOAEs} up to 4.5 kHz ($R^2 = 0.02$; $p = 0.21$) only. In the NTL group, however, a weak negative correlation between emission level and frequency was found ($R^2 = 0.27$; $p < 0.000$) group. In both language groups the SOAE width was negatively correlated with the SOAE level (Fig. 3.2B). Thus, large emission peaks were significantly narrower than SOAEs with smaller levels (using ANOVA) in the TL ($R^2 = 0.13$; $p < 0.001$) and NTL ($R^2 = 0.12$; $p = 0.001$) group.

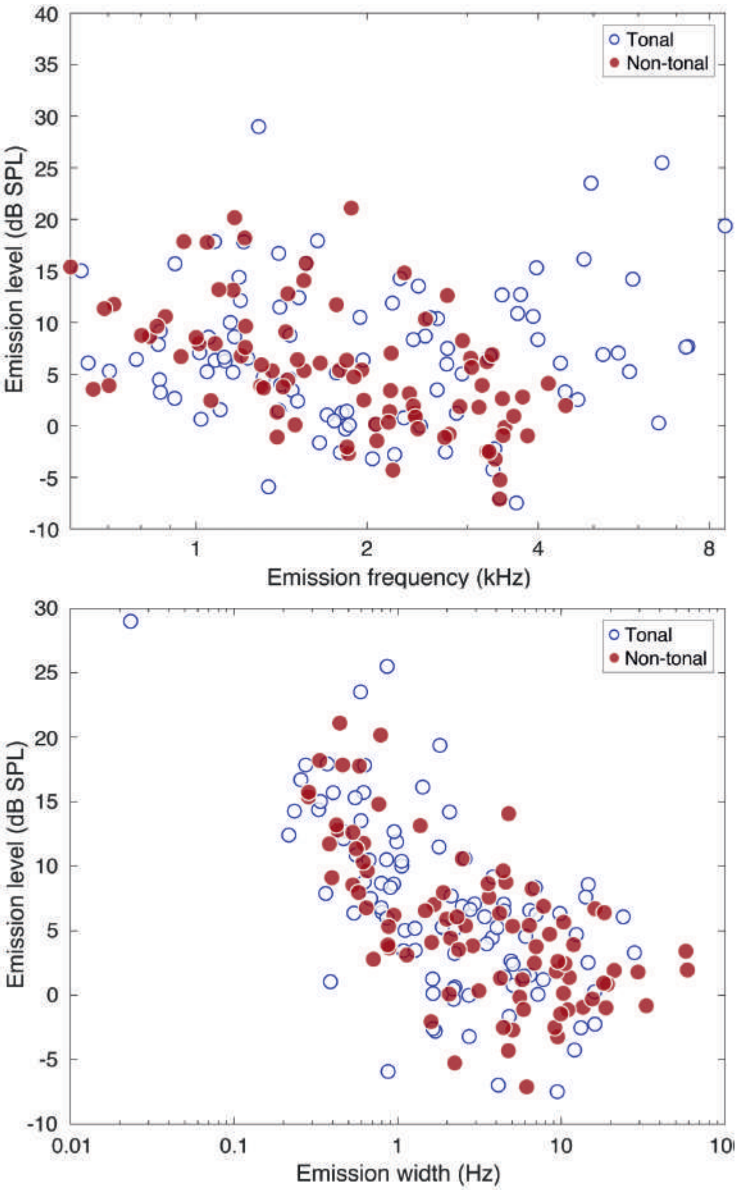


Figure 3.2. Characteristics of unsuppressed SOAEs. In both panels open circles indicate data of TL and filled circles that of NTL native speakers. (A) The relationship between SOAE frequency and level. (B) The relationship between SOAE level and width. In both language groups the SOAE width was significantly negatively correlated with the SOAE level (TL: $p < 0.001$, $R2 = 0.13$ and NTL: $p = 0.001$, $R2 = 0.12$).

In summary, for both language groups, we found the well-known difference that females show more SOAEs than males. Compared to the NTL native speakers, TL native speakers had better hearing thresholds and more SOAEs at frequencies between 4.5 and 8 kHz.

3.3.2 Suppression tuning curves (STCs)

Here, STCs of TL and NTL native speakers were compared, to evaluate whether increased frequency selectivity could also be observed at cochlear level. STCs were asymmetrically V-shaped and selectively tuned (Fig. 3.3). We evaluated the slopes of both STC flanks of the TL native speakers (n = 70) and the NTL native speakers (n = 69). Average STC slopes did not differ significantly between groups. The average low frequency slopes were 35 and 41 dB/oct, and the average high frequency slopes were 46 and 45 dB/oct, respectively, for the TL and NTL group.

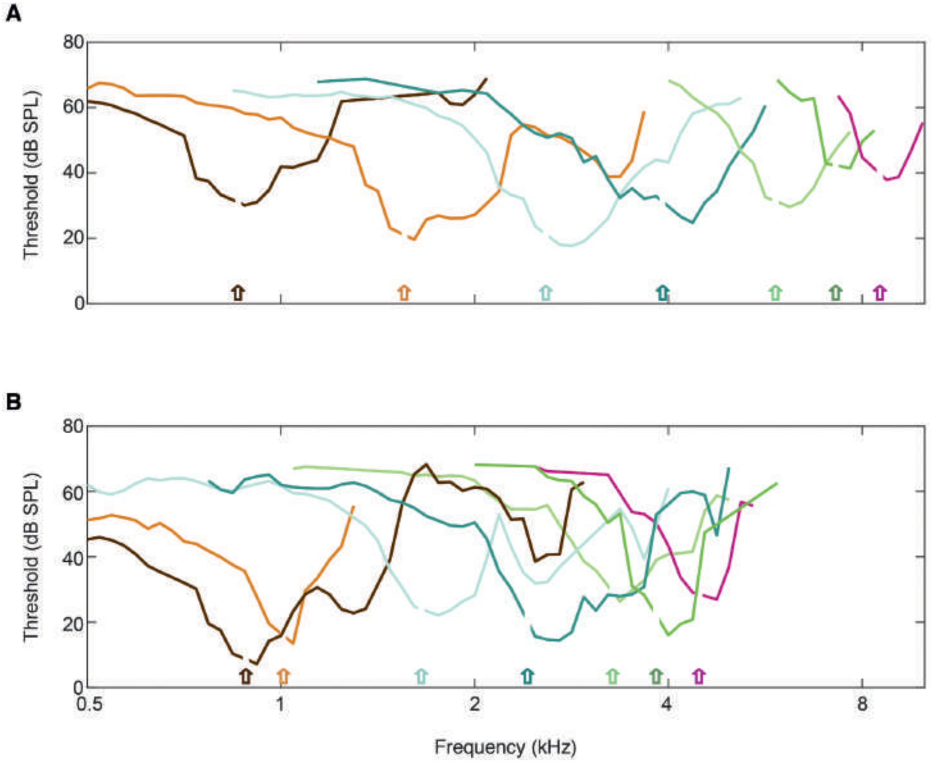


Figure 3.3. Representative STCs of SOAEs from different TL (A) and NTL native speakers (B). STCs indicate the stimulus level needed for 3 dB suppression of the SOAE. The arrows indicate the SOAE frequencies. The colors match the corresponding STC. The stimulus frequencies within 10 Hz of the unsuppressed SOAE frequency were omitted (see main text) and appear as gaps in the STC. Note that in both language groups some STCs contain secondary suppression lobes.

The f_{tip} (most sensitive frequency) of the STC could fall on either side of the emission frequency, but was typically above the emission frequency. The f_{tip} in TL native speakers was on average 5.8% above the unsuppressed SOAE frequency, versus 4.3% in the NTL group. The level at f_{tip} (STC's best threshold) between both native language groups differed in their median with 3 dB, median thresholds being 22.1 dB SPL for the TL native speakers and 19.0 dB SPL for the NTL native speakers. We were interested whether SOAE level correlates with suppression threshold (Fig. 3.4). Interestingly, the suppression thresholds in TL native speakers were independent from SOAE level, whereas the suppression threshold was significantly negatively correlated to emission level in the NTL group ($p < 0.001$). When evaluating the frequency range up to 4.5 kHz only, the suppression threshold remained independent from the SOAE level in the TL group.

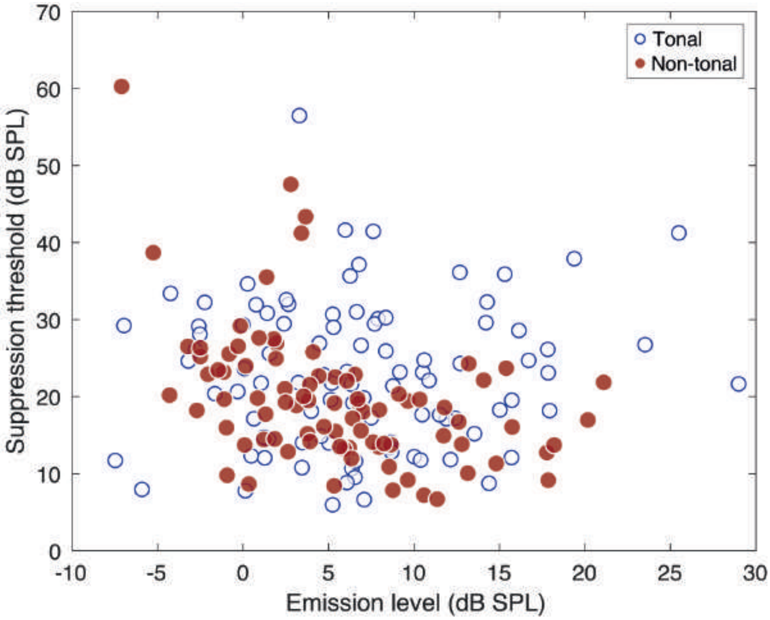


Figure 3.4. The suppression thresholds of 3 dB STCs in relation to SOAE level. Open circles indicate data of TL and filled circles that of NTL native speakers. In the TL group the suppression threshold was not correlated with emission level, whereas in the NTL group both were significantly ($p < 0.001$) correlated.

Tuning curve sharpness and tuning curve side-lobes

STCs showed the typical asymmetric shape, with steeper slopes for the high-frequency flank. In Figure 3.5A we show the average STC per language group, that represents at least 10% of the data, with standard deviation. All STCs were aligned with respect to the tip frequency and level. The averaged STCs were very similar between both language groups. The frequency selectivity was defined as the filter quality factor Q_{10dB}, for all subjects (Fig. 3.5B). Tuning was similar in both language groups (median Q_{10dB} TL: 4.28; median Q_{10dB} NTL: 4.81) and independent from f_{tip}.

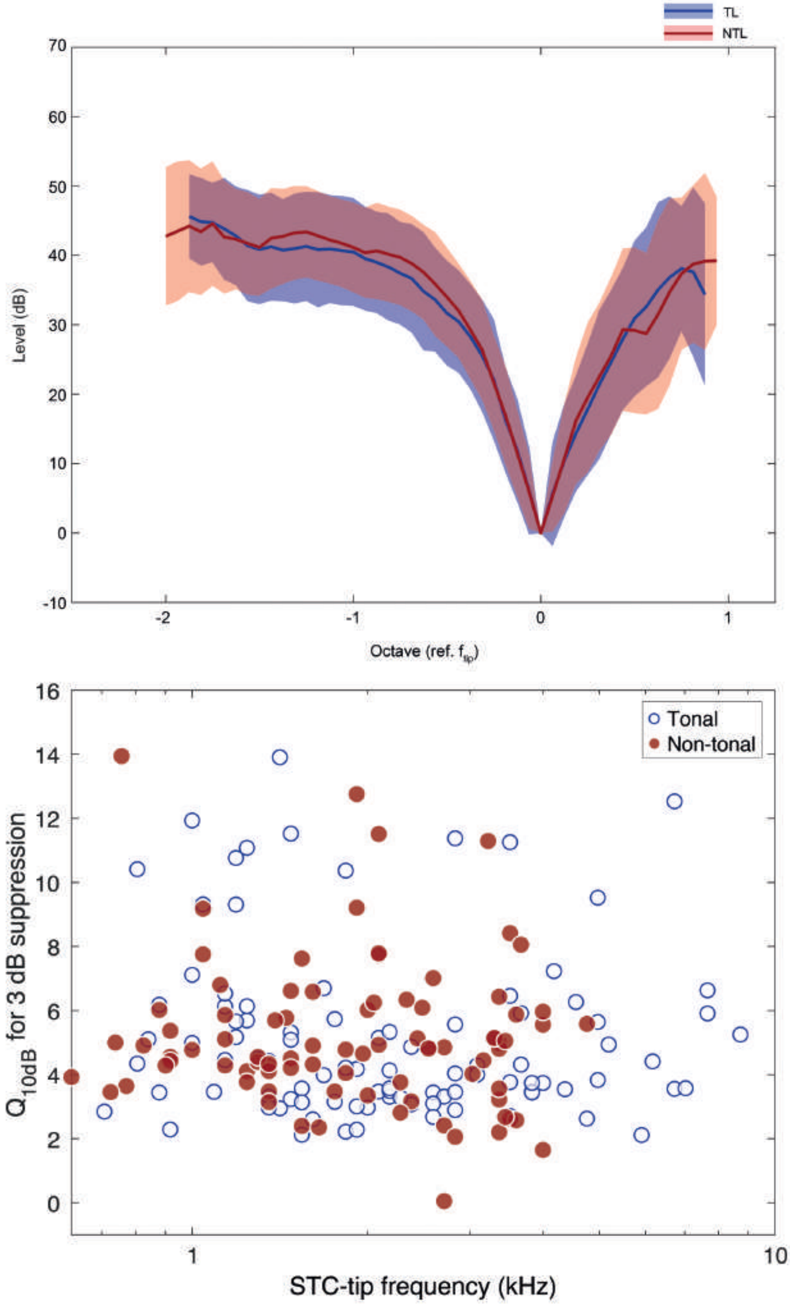


Figure 3.5. Comparison of frequency selectivity between both language groups. (A) The average STC per language group with standard deviation (sd). Data shown represents at least 10% of the STCs. (B) Comparison between filter quality measures between TL (open symbols) and NTL native speakers (filled symbols). The filter quality factor Q_{10dB} was determined from STCs. Tuning selectivity in both language groups was similar and did not correlate with STC_{tip} frequency.

Side-lobes represent an additional suppression area at the higher frequency flank of the STC, in some cases even two side-lobes could be observed (Fig. 3.3). In Figure 3.6 we show the frequency of the side-lobes, as a function of frequency of the main STC-tip. Primary side-lobes were in general 0.5-1 octave above the emission frequency. We observed STCs with primary side-lobes in 38% of the emissions for the TL group and in 37% of the NTL group. Additional secondary side-lobes were recorded less commonly. Of the STCs with primary side-lobes, the TL native speakers rarely had secondary side-lobes (11%), whereas secondary side-lobes in the NTL group were recorded more frequently (38%).

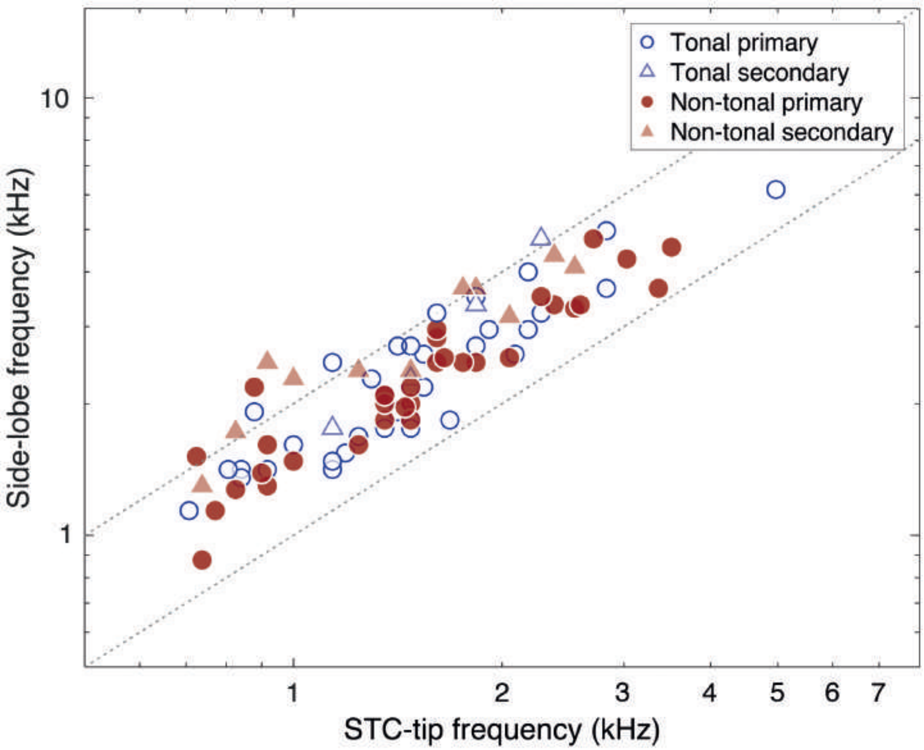


Figure 3.6. Frequencies of the primary and secondary side-lobes of STCs, as a function of the frequency of the primary STC-tip. The dashed diagonal lines are added for orientation and are 1 octave apart. Values of the side-lobes of TL (open circles) and NTL native speakers (filled circles) and secondary side-lobes of TL (open triangles) and NTL (filled triangles) are shown. In both language groups the average ratio between the STC-tip and the first side-lobe is 1:1.5 (sd TL: 0.3; sd NTL: 0.3). The average ratio between the STC-tip and the second side-lobe is in the TL group 1:1.8 (sd: 0.3) and in the NTL group 1:2.0 (sd: 0.3).

3.4 Discussion

The properties of spontaneous otoacoustic emissions (SOAEs) were compared between native speakers of a tonal language (TL) and those of a non-tonal language (NTL). SOAEs of both language groups were similar in all aspects we investigated, except for the range of frequencies at which SOAE were detected. In the TL group, emissions were detected above 4.5 kHz, whereas SOAE frequencies in the NTL group never exceeded this frequency. We specifically evaluated frequency tuning curves of SOAE suppression, because of the possible role of cochlear frequency selectivity in language processing. However, we found no difference in frequency selectivity between both language groups.

Our findings correspond with previous research that reported that Asians are more likely to emit SOAEs at higher frequencies compared to Caucasians (Whitehead et al., 1993; Chan and McPherson, 2001). The occurrence of high frequency SOAEs can potentially be caused by (1) middle ear and (2) inner ear differences between both groups. In general, a shorter meatus, smaller middle ear canal volume, and a smaller tympanic membrane, increase the high frequency transmission (Plassmann and Brändle, 1992). Models have also shown that such changes in middle ear mechanics influence otoacoustic emissions (Avan et al., 2000). On average, Asians have smaller ear canal volumes when compared to Caucasian subjects (Whitehead et al., 1993; Chan and McPherson, 2001; Wan and Wong, 2002; Shahnaz and Davies, 2006). Asians who emitted SOAEs at higher frequencies in fact had smaller ear canal volumes and static admittance than Caucasians (Chan and McPherson, 2001). Consequently, the middle ear characteristics of Asians may favor the transmission of high frequencies. This would not only affect the SOAE transmission towards the outside but also the transmission of high frequency sounds into the ear, which could explain lower hearing thresholds at higher frequencies in TL native speakers (Fig. 3.1A).

Recently, peripheral frequency selectivity has been investigated using a number of measures. All these measures provide an estimate of the quality factor of cochlear filters, either expressed as Q_{10dB} or Q_{ERB}. In general, measures which are believed to be unaffected by cochlear nonlinearity (compression), provide relatively high values for the quality factor of cochlear filters. In humans, Q-values obtained using these methods range from about 15 to 20, where larger Q-values are measured for cochlear filters with higher center frequency. These linear methods include measurements of SFOAE group delays (Shera et al., 2002) and forward masking psychoacoustic tuning curves (Shera et al., 2002; Oxenham and Shera, 2003). In contrast, measurements of peripheral frequency selectivity, that depend on cochlear compression, provide broader filter estimates with Q-values of approximately 4-5. These methods include measurements of suppression tuning curves of otoacoustic emissions (SOAEs: Zizz and Glattke, 1988; Manley and van Dijk, 2016; SFOAEs: Charaziak and Siegel, 2014; DPOAEs: Abdala et al., 2007) and psychoacoustic tuning curves derived by simultaneous masking (Moore, 1978; Oxenham and Shera, 2003). Smaller Q-values are presumably based on the compressive action of the mechanical response of the basilar membrane, which shows broader spatial excitation patterns at higher sound levels (Robles and Ruggero, 2001).

Notably, SOAE suppression, as studied in the current paper, must inherently depend on cochlear compression. In general, when two signals are processed by a compressive nonlinearity, the stronger signal determines the degree of compression, which then affects the smaller signal more than if it was present alone. The smaller signal is thus suppressed. Hence, when the external tone interacts with the SOAE, the latter will be suppressed when the excitation due to the tone becomes larger than the excitation of the SOAE. Models suggest that the vibration pattern of an SOAE is maximal near the tonotopic place corresponding to the emission frequency (Epp et al., 2015; for an animation see the supplemental material of Manley and van Dijk, 2016). Consequently, it can be assumed that SOAE suppression by an external tone reflects tonotopy and frequency selectivity of the basilar membrane.

In short, measures of cochlear tuning that are believed to reflect near-threshold linear cochlear behavior, provide highly selective estimates of cochlear tuning. There is substantial evidence that the Q-values of these linear measurements correspond to those of auditory neurons (Shera et al., 2002; Joris et al., 2011; Sumner et al., 2018), although the comparison includes the assumption of a factor, referred to as the "tuning ratio". In contrast, in measurements based on experimental protocols that presumably engage cochlear compression, estimates of Q-values are lower. These nonlinear measurements suggest broader tuning, which reflects the fact that cochlear mechanical excitation patterns are wider at higher sound levels. Consistently, the Q-values of SOAE suppression tuning are below that of neural tuning curves in mammals (macaque: cf. Martin et al., 1988 with Joris et al., 2011) and also birds (barn owl: cf. Engler et al., 2020b with Köppl, 1997a, b). In the present study, we used the nonlinear measurement paradigm of SOAE suppression tuning. As described above, this measurement presumably reflects cochlear frequency selectivity, although it may not be a direct measure of nerve fiber tuning. Nevertheless, differences in mechanical cochlear tuning between TL and NTL participants likely would have been detected if they existed. Hence, our results suggest that there is no difference in tuning of the basilar membrane between TL and NTL speakers.

The only aspect where STC were different between TL and NTL speakers was the number of secondary side-lobes that were observed. In TL speakers, these side-lobes were more common than in NTL speakers. At present we can only speculate about an explanation for this difference. Spontaneous otoacoustic emissions are believed to correspond to standing waves in the cochlea (Kemp, 1980; Shera, 2003; Epp et al., 2015).

In a model of basilar membrane mechanics, these standing waves have antinodes at well-defined positions along the basilar membrane (Epp et al., 2015). Manley and van Dijk (2016) suggested that the tonotopic frequency of these antinode positions correspond with side-lobes in the STC. Thus, the side-lobes may be a consequence of interactions between the external tones in the STC measurements and the antinodes of the standing wave. Note, that the standing wave occurs between the stapes footplate of the middle ear and the tonotopic location of the SOAE frequency. Possibly, the differences between side-lobe properties found in the current study may reflect subtle differences in the mechanical properties of the middle ear, as described earlier in this section. However, at present this remains speculative.

Behavioral studies have shown that Asian TL native speakers outperform Caucasian NTL native speakers in pitch perception accuracy (e.g.: Deutsch et al., 2004; Pfordresher and Brown, 2009; Hove et al., 2010; Braun and Johnson, 2011; Giuliano et al, 2011). Our aim was to evaluate whether this enhanced pitch perception of TL speakers reflects sharper frequency selectivity at cochlear level. As a measure of cochlear frequency selectivity, we evaluated STCs of SOAEs. Possible mechanisms behind the enhanced pitch perception of TL native speakers, and to what extent it has a basis in cochlear frequency selectivity, is discussed in the following paragraphs.

The absence of a difference in cochlear tuning suggests that more central structures are responsible for the better pitch acuity in TL speakers. Speech processing is mainly mediated in the left-brain hemisphere, the area where also the temporal information is encoded (Zatorre et al., 2002). Studies have shown that human speech understanding is primarily achieved by temporal processing rather than frequency selectivity (Shannon et al., 1995). For TL native speakers the detection of time varying pitch contours is essential for their native language understanding. At the cochlear level, wider auditory filters would theoretically lead to an improvement of temporal processing. However, wider auditory filters would cause poorer spatial resolution which consequently results in a decrease of frequency selectivity. We did not detect such a significant difference in Q_{10dB} for TL native speakers. Therefore, there was neither evidence for enhanced frequency selectivity nor for better temporal processing at cochlear level of Asian TL speakers, that would explain their behavioral outperformance in pitch perception.

Acoustical training is linked to the development of enhanced pitch perception. This training effect appears to generalize across linguistic and non-linguistic specific contexts. Musicians, for example, do not only detect frequency-movements of pure tones very precisely, but also perceive pitch-contours in linguistic manipulations more accurately (e.g.: musician children: Magne et al., 2006; professional musicians: Schön et al., 2004). Therefore, musical training seems to favor the processing of linguistic relevant pitch information in Mandarin Chinese (Wong et al., 2007). Moreover, TL acquisition (as a form of acoustical training) leads to pitch accuracy in non-linguistic contexts as well (e.g. Salmelin et al., 1999; Vihla et al., 2002; Deutsch et al., 2004; 2006; Pfordresher and Brown, 2009). In this study we included non-professional musicians only (definition of professional musician: Micheyl et al., 2006), thus any possible training effect would be primarily related to the native language and therefore stable within each group.

Acoustical experience favors the accurate behavioral perception of tones, as the TL acquisition is linked to an increased accuracy of tone perception (Deutsch et al., 2004; 2006; Pfordresher and Brown, 2009). We speculate that acoustical training, due to language acquisition, enhances the pitch perception abilities of TL native speakers. Studies have addressed a link between TL experience and enhanced pitch representation and tracking. However, this enhancement could not be fully explained by increased temporal pitch processing (Krishnan et al., 2005). Therefore, it was hypothesized that language experience induces neural plasticity at the brainstem level. In fact, TL native speakers showed enhanced pitch encoding measured at the brainstem (Krishnan et al., 2005) and cortical pathways (Kuhl, 2004). In other words, language experience affects the neural pathways at subcortical brainstem level and the central level of the auditory cortex.

Moreover, behavioral and imaging studies have shown that speech processing networks develop which are language-specific (e.g. Gandour et al., 2000; Zatorre et al., 2002). When infants learn their native language, their brains develop language specific networks (Kuhl, 2004; Krishnan et al., 2010b). TL native speakers seem to use different neural networks depending on whether the change in pitch is linguistically relevant or not (Gandour et al., 1998; Wong et al., 2004; Pfordresher and Brown, 2009). Thus, the sensitivity is increased to sounds that are similar to those of this particular language (Kraus and Banai, 2007; Krishnan et al., 2010a). Consequently, the auditory system appears to be experience-based and plastic in modification. Experience dependent neural ascending and descending pathways optimize the functionality and form the auditory cortex (Suga et al., 2003). Interestingly, such processing pathways are not strictly restricted to language specific cues (e.g. Bent et al., 2006). TL acquisition tunes the overall neuronal response to pitch in the brainstem with enhanced sensitivity to speech relevant cues (e.g.: Swaminathan et al., 2008; Krishnan et al., 2009). Thus, effects of acoustical training can generalize across linguistic and non-linguistic specific contexts (e.g.: Bidelman et al., 2013). Presumably, this is how TL acquisition (as a form of acoustical training) can lead to pitch accuracy in non-linguistic context as well (e.g. Salmelin et al., 1999; Vihla et al., 2002; Deutsch et al., 2004; 2006; Bent et al., 2006; Pfordresher and Brown, 2009).

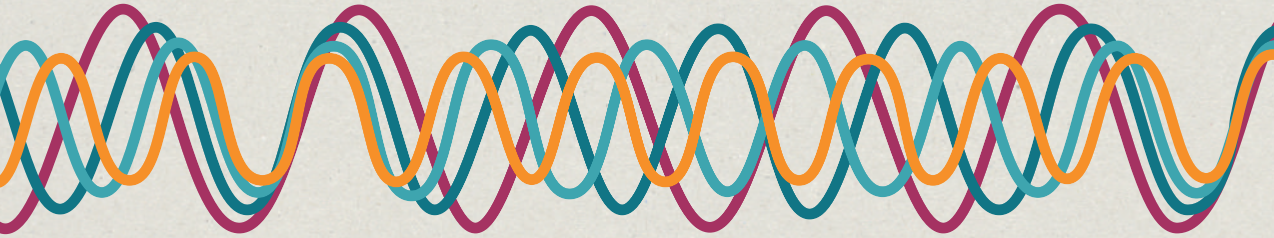
In addition to differences in acoustical training and exposure, there are also differences in gene expression associated with pitch perception. Specific genes may be linked to enhanced pitch perception (Zatorre, 2003; Schellenberg and Trehub, 2008; Hove et al., 2010). This aspect becomes especially important when testing whether native language experience can be ruled out as a training factor for pitch perception, for example when testing Asians that grew up without exposure to a TL.

3.5 Conclusions

In this study, SOAEs of Asians with TL acquisition and Caucasians with no TL experience were recorded and suppressed by pure-tone stimulation. Suppression tuning curves were similar between both language groups. This suggests that the enhanced frequency selectivity of Asian TL native speakers is not based on a difference in cochlear processing. SOAEs above 4.5 kHz were found in TL native speakers only, which is probably based on differences in middle ear properties.

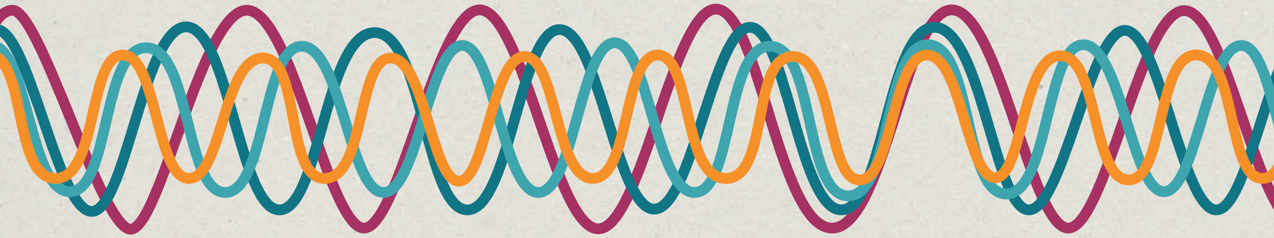
Acknowledgments

This work was supported by the EGRET (European Glaucoma Research Training Program, Marie Sklodowska-Curie-COFUND grant; No 661883). A special thanks to Paolo Toffanin for programming.



Chapter 4

Relationship between irregularities in spontaneous otoacoustic emissions suppression and psychophysical tuning curves



S. Engler, E. Gaudrain, E. de Kleine, P. van Dijk

This chapter has been published as Engler et al., 2022.
The Journal of the Acoustical Society of America 151 (2), 1055-1063.

Abstract

The suppression of spontaneous otoacoustic emissions (SOAEs) allows the objective evaluation of cochlear frequency selectivity by determining the suppression tuning curve (STC). Interestingly, some STCs have additional side-lobes at the high frequency flank, which are thought to result from interaction between the probe tone and the cochlear standing wave corresponding to the SOAE being suppressed. Side-lobes are often in regions of other neighboring SOAEs but can also occur in the absence of any other SOAE. The aim of this study was to compare STCs and psychoacoustic tuning curves (PTCs). Therefore, STCs and PTCs were measured in: (1) subjects in which the STC had a side-lobe, and (2) subjects without STC side-lobes. Additionally, PTCs were measured in subjects without SOAEs. Across participant groups, the quality factor Q_{10dB} of the PTCs was similar, independently from whether SOAEs were present or absent. Thus, the presence of an SOAE does not provide enhanced frequency selectivity at the emission frequency. Moreover, both PTC and STC show irregularities, but these are not related in a straightforward way. This suggests that different mechanisms cause these irregularities.

4.1 Introduction

Frequency selectivity is the ability to separate a complex sound into its pure tone components (e.g., Moore, 1989). The auditory frequency selectivity of an individual can be determined through objective, e.g., suppression tuning curves (STCs) or behavioral measurements, e.g., psychophysical tuning curves (PTCs). Physical measurements of frequency selectivity (STCs) rely on recordings of physiological properties from the auditory system, whereas behavioral measurements (PTCs) rely on behavioral responses from the participants. Both measures are based on interaction between signals. In one type of STC, spontaneous otoacoustic emissions (SOAEs) are suppressed by external tones. This suppression is thought to be based on overlap between the basilar membrane (BM) excitation patterns of the SOAE and that of the suppressive tone. In simultaneous PTCs, the perception of a target tone interferes with the presentation of a masker, for example, a narrow band noise. The perception of the target tone can be hindered when the BM excitation pattern of the tone overlaps with that of the masker. For STCs, frequency selectivity is determined by measuring the level of the suppression tone that yields a certain level of SOAE suppression, for different frequencies of the tone. Similarly, for PTCs, frequency selectivity is determined by measuring the masker level that corresponds to the detection threshold of the target tone, as a function of the masker frequency. The tuning curves (TCs) are essentially V-shaped, where the tip of the tuning curve, the most sensitive point, is close to the SOAE in STCs, or the target tone in PTCs.

Although the methods used to obtain STCs and PTCs have a lot in common, the two measures reflect very different processes. While PTCs require the listener to solve an auditory scene analysis problem and make a decision about the presence or absence of a tone embedded in noise—thus, involving the whole peripheral and cortical auditory system—STCs are limited to the most peripheral level. Understanding the differences between the two measures can thus potentially reveal how frequency selectivity evolves from the periphery to more central brain areas (Langers and van Dijk, 2012; Moerel et al., 2012). Further, comparing PTCs (which can be measured in virtually any normal-hearing listener) to STCs (which can only be measured in listeners with SOAEs) offers the opportunity to better understand the mechanisms underpinning the presence of SOAEs, and what effect they may have on perception.

When comparing frequency tuning measured with SOAE-STCs and simultaneous masking PTCs, both measures are in good alignment across individuals (Zizz and Glattke, 1988). Still, it was suggested that the presence of SOAEs influences the appearance of PTCs. According to the literature, spontaneously emitting ears appear to have sharper PTCs when using simultaneous, tonal maskers, at least at, or close to, the SOAE frequency (Bright, 1985; Micheyl and Collet, 1994). It was suggested that PTCs in which the target tone was centered to an SOAE were more sharply tuned compared to ears without SOAEs (Bright, 1985). Baiduc and colleagues (Baiduc et al., 2014) also investigated the relation between PTC sharpness in ears with and without SOAEs. However, they used narrowband noise maskers rather than tones, and found no difference in frequency tuning at emission frequency. The interaction between a tonal masker and an SOAE can be perceived by the participant as beating (Wilson, 1980; Long and Tubis, 1988a), which could explain the sharper PTCs observed at emission frequency by Bright (1985). To avoid the generation of such perceptual cues, we chose a simultaneous noise masker instead.

Because TCs are generally V-shaped and reflect selectivity, they are primarily characterized by their sharpness. However, some TCs show irregularities, causing them to deviate from the typical V-shape. In SOAE-STCs, additional suppression side-lobes were observed, characteristically at about 0.5 and 1 octave above the SOAE frequency (Manley and van Dijk, 2016; Engler et al., 2020a). It has been suggested that these side-lobes reflect interactions between a BM standing wave corresponding to the emission frequency and the presented tone (Manley and van Dijk, 2016). Alternatively, interactions between multiple SOAEs in the same ear could result in irregular STCs. PTCs can also deviate from the V-shape, with peaks and dips at frequencies away from the probe-tone frequency (Baiduc et al., 2014). These irregularly shaped PTCs appear to be more common in ears with SOAEs (Baiduc et al., 2014), suggesting that the presence of SOAEs could influence the perception of target tones. In PTCs, the masker could produce similar dips, as seen in SOAE-STCs produced by the presented tone.

The objective of the present study is to compare STCs to PTCs, both in their sharpness and in the presence or absence of irregularities in their shape. Since the irregularities of both the STCs and PTCs have been attributed to similar cochlear mechanisms, we hypothesized that PTCs measured with a target tone placed at an SOAE frequency would yield TC irregularities that are related to side-lobes in the STC of that emission. Conversely, if the STC does not show side-lobes, the corresponding PTC would be expected to be simply V-shaped. To verify this, STCs and PTCs were measured in participants with SOAEs. Both types of STCs, namely with and without side-lobes, were included. Additionally, PTCs were also obtained from listeners without detectable SOAEs. In the absence of SOAEs, no irregularities in the PTCs are expected.

The aim of this study was to directly compare the frequency selectivity and the shape of SOAE-STCs to PTCs from the same ear, tested at the SOAE frequency. We investigated whether the additional side-lobes in STCs are reflected in the shape of PTCs and whether PTC side-lobes are absent in people without SOAEs.

4.2 Material and methods

4.2.1 Participants

Prospective participants[1] were screened, in both ears, for the presence of SOAEs and for normal hearing. Only ears that showed clear SOAE peaks of at least 3 dB above the microphonic noise, or ears without any recordable SOAE peaks, were chosen. After screening, 26 participants were included. One ear per participant was included in the measurement. The chosen ear was then categorized in one group with SOAEs (n = 17) and one group without SOAEs (No SOAE group, n = 9). The group with SOAEs was subdivided depending on whether STC side-lobes were present (SL group, n = 8) or absent (NSL group, n = 9). A high frequency STC side-lobe was defined as the reduction in STC-slope in relation to the expected high frequency slope, followed by a relatively abrupt increase in slope again. Thus, side-lobes were visible as secondary dips in the STC. The number of SOAEs per tested ear did not differ significantly (Mann–Whitney U = 43.5, p = 0.48) between the SL and NSL group. However, participants with eight or more SOAEs were observed in the SL group only.

Participants self-reported no history of ear pathologies. All participants had normal hearing with pure tone thresholds 25 dB hearing level at octave frequencies between 0.25 and 8 kHz (Audiosmart, Echodia, Clermont-Ferrand, France). The median age of all tested adults (n = 26) was 29 years (ranging from 22 to 48 years). In total, 11 participants had some degree of musical training. With the highest level of musical training, seven of these can be seen as non-professional musicians. None of the included participants had been playing an instrument for more than 10 years and none were practicing music daily (definition of professional musicians, Micheyl et al., 2006). All participants were naive to the test-routine. Nine subjects participated in forced choice experiments before, but none in a similar masking study with simultaneous noise and a three-alternative forced choice routine.

The majority of the included participants with SOAEs were female (82.4%). Participants with STC side-lobes were exclusively female (n = 8), whereas STCs without side-lobes (n = 9) were also recorded in males (33.3%). In the participant group without SOAEs, the sex representation was more balanced, with slightly more males than females (55.6%).

4.2.2 Measurements

All measurements were carried out in a soundproof anechoic room of the ear nose throat (ENT) department within the University Medical Centre Groningen (UMCG, Groningen, Netherlands).

SOAE recordings

To record the SOAEs, an occluding soft foam ear plug, which included the Etymotic ER-10B microphone-speaker system (Etymotic Research, Inc., Elk Grove Village, IL), was placed in the participants' external ear canal. For the STC recording, an amplification of the microphone output of 60 dB was applied, by adding the 40 dB gain of the Etymotic ER-10B system (Etymotic Research, Inc., Elk Grove Village, IL) and a 20 dB gain by the

Stanford Research Systems amplifier (SRS Inc., model SR 640, Sunnyvale, CA). The microphone signal was monitored using a spectrum analyzer (SRS Inc., model SR 760, Sunnyvale, CA). A MOTU 624 soundcard (MOTU Inc., Cambridge, MA) was used to record the microphone signal and to generate the tone stimuli. The playback of stimuli and recording of SOAEs was controlled by custom routines developed in Matlab (MathWorks Inc., 2016a, Natick, MA) using a 24-bit resolution and a 48 kHz sampling rate.

A Lorentzian function was fitted to emission recordings with the best signal-to-noise ratio to estimate the SOAE characteristics, such as emission frequency (f_{SOAE}), width, and amplitude.

STC measurements

The STCs were obtained following a procedure adapted from that described by Engler et al. (2020a). The procedure assesses the suppression of SOAEs by a range of subsequent stimulus tones. Stimulus tones were generated using an ER-2 driver connected to the ER-10B microphone. In an automated procedure that lasted approximately 1 h, 1.2 s probe tones, with 10 ms raised-cosine ramps, were presented ranging over 70 frequencies (0.5 – 9.9 kHz, in 1/16 octave steps) and 24 levels (0–70 dB SPL, in 3 dB steps), in quasi-random order. Tone levels were calibrated in situ for each participant, with the emission probe placed in the ear canal, and using the emission microphone for calibration. For each presented stimulus, the SOAE level was computed. A tonal signal with a frequency equal to the stimulus (f) plus two higher harmonics (2f, 3f) was fitted to the recorded signal by a least squares minimization procedure. The fitted stimulus was then subtracted from the recorded signal to include the SOAE but not the presented stimulus. SOAEs were filtered by a narrow fast Fourier transform (FFT) bandpass filter to estimate the SOAE-peak frequency. For details of the STC analysis, the reader is referred to Engler et al. (2020a). The reference magnitude of the emission was estimated using the Hilbert transform as the mean of the module of the analytic signal expressed in decibels. The suppression was calculated as the difference between the measured emission level at a given probe tone level and frequency and the reference emission level.

The STC was calculated by estimating the 3 dB isocontour using Matlab's contour function (Fig. 4.1). Multiplencontours can be returned, as local noise in the recordings can cause accidental closed contours to arise in areas remote from the STC. The obtained contours were thus curated with an automatic routine to keep only the contour depicting the STC.

PTC measurements

The PTCs were measured by estimating the masking threshold of a target tone by narrow bands of noise centered on different probe frequencies covering the expected shape of the TC. Because potential side-lobes were expected to fall within 0.5 to 1 octave above the PTC-tip, as seen in STCs (Manley and van Dijk, 2016; Engler et al., 2020a), the probe frequencies ranged from 0.9 to þ1.4 octaves relative to the tone frequency, by steps of 0.1 octave.

For each probe frequency, the threshold was obtained using an adaptive (2-down, 1-up) three-interval, three-alternative forced choice (3I-3AFC) method. Two of the intervals were the masker alone, while the odd-one-out, which the participants were instructed to detect, contained the target tone and the masker. Participants were shown an interactive interface that displayed three buttons on a computer screen. Each button lit up when an interval was presented. The participant then had to choose the button corresponding to the odd-one-out and was given visual feedback after responding.

When two correct responses were given in a row, the level of the masker was increased by the step size, and for every incorrect response, the level was decreased by the step size. The step size started at 5 dB but was decreased to 2 dB after the first reversal. The procedure continued until 8 reversals were observed; the masking threshold was calculated as the average of the masker level over the last 6 reversals.

The target tone was centered on the emission frequency of interest, and lasted 200 ms with 20 ms raised cosine ramps. The masker had a duration of 300 ms, with 20 ms raised cosine ramps, and consisted of a band of noise of 1/3 octave width (slope 288 dB/octave) centered on the probe frequency. When the target and masker were added together, the target was temporally centered in the masker. The intervals were separated by a silence of 300 ms. The target level was set to 10 dB sensation level, after the absolute threshold was measured. The initial masker level was adapted to the probe frequency, following the general shape of PTCs reported in the literature (adjusted from Moore, 1978) in order to reduce unnecessary trials and arrive close to the threshold.

When hovering close to the detection threshold for too long, the target tone detection gradually became very difficult. To help the participant, a “target reminder” was presented after four consecutive incorrect responses, as well as at the beginning of each adaptive track, i.e., every time a new masker frequency started. Participants did not need to respond to this target reminder.

During the PTC measurement, stimuli were generated by two ER-2 devices which were connected to the ER-10B emission probe that was inserted into the ear canal. The same settings were used as for the STC measurements. Except for two participants, the masking-noise output was presented through the headphone driver (HB 7 Headphone Buffer, Tucker-Davis Technologies, Alachua, FL).

4.2.3 Procedure

The STC and PTC recording procedures encompassed three main steps: (1) 2 min SOAE recording without any external stimuli. For participants with SOAEs, this step was followed by (2a) the STC measurement (± 60 min). Further, for all participants, (2b) the PTC were then measured (60 min). Finally, (3) a 2 min SOAE recording, equivalent to step 1, was repeated to ensure SOAEs had not changed.

The STC data were (except for one participant) collected prior to the PTC measurement. For the majority of participants, the STC measurements were taken from an earlier experiment; thus, the data are partly included in the publication of Engler et al. (2020a). In all cases, steps 1 and 3 were performed twice (for both sessions). Comparing the SOAEs recorded on both measurement days, they were in good agreement with each other. Median differences between STC and PTC measurement day were 1.59 Hz (standard deviation, sd: 9.09 Hz) for SOAE frequencies and 0.36 dB (sd: 4.47 dB) for the SOAE levels.

For participants with SOAEs, the PTCs were measured at the selected SOAE frequency. For the participants that did not have SOAEs (No SOAE group), the PTCs were measured at a frequency that matched that of a participant with SOAE. Details of the participant grouping are provided in Supplementary Table I (see supplementary material[2]).

To make sure that the three-alternative forced choice task was well understood from the beginning of the PTC measurement, all participants received the same training prior to the actual measurement. During the training, the frequency of the target tone was always set to 2.5 kHz. The absolute threshold determination of the target tone was followed by a shortened procedure where only masker frequencies at 2.18 and 2.87 kHz were tested.

4.2.4 Data analysis

For both STC and PTC, the tip was defined as the lowest point in the vicinity of the target frequency. The frequency selectivity can be quantified as the filter quality factor Q_{10dB} of the TC. This factor is defined as the ratio between the tip-frequency and the width of the TC at 10 dB above its tip, using linear interpolation, if necessary.

Another way of characterizing sharpness is to estimate the slope of the TC for each flank. This has the potential of revealing the asymmetries in TC shapes, whereas the Q_{10dB} only concerns the width. The slopes of the low- and high-frequency flank were estimated as the average slope, obtained by linear interpolation, on a segment of the TC between the tip and 25 dB above it. The average TCs were calculated with a generalized additive model (GAM) based on likelihood-based fits, presenting confidence intervals of 95%.

Statistical analyses were performed in IBM SPSS Statistics (Version 23, Armonk, NY) and Matlab. Because of the small sample size, we could not guarantee that the data represent a normal distribution; therefore, we chose a non-parametric test. The Kruskal–Wallis test was used for comparisons across tested groups, reporting the H statistic. For pairwise comparisons between STC and PTC recordings within and across the tested groups of participants, the Mann–Whitney U test was applied, also reporting the H-statistic.

Smooth average estimates of the TCs for each group were obtained in R (v4.0.3, R Foundation for Statistical Computing, Vienna, Austria) using a generalized additive model (mgcv, v1.8–33, Wood, 2011) as implemented in ggplot2 (v3.3.3, Wickham, 2016).

4.3 Results

Data extracted from the TCs to compare them were TC-tip level and tip-frequency, the high and low TC-slopes, and the Q_{10dB}. Overall STCs and PTCs showed the typical V-shape, with asymmetric flanks. TC-tips were near the SOAE frequency in STCs or the target tone in PTCs. A selection of representative TCs is displayed in Fig. 4.1. Note that STC side-lobes may occur in absence (Fig. 4.1A or presence (Fig. 4.1B and C) of other neighboring SOAEs. Corresponding PTCs may also show irregularities in tuning curve shape (Fig. 4.1A and B), but may also be smoothly V-shaped (Fig. 4.1C). PTCs may have side-lobe-like irregularities in the absence of STC side-lobes or neighboring SOAEs (Fig. 4.1D), as well as in the presence of neighboring SOAEs (Fig. 4.1B). Irregular and smoothly shaped PTCs occur also in participants without SOAEs (Fig. 4.1F and G).

Figure 4.2 shows all individual tuning curves, as well as the averaged tuning curves per subject group. STCs of the SL group showed a clear difference at the high frequency flank compared to the NSL group (Fig. 4.2A). The STCs of the SL and the NSL group deviate significantly (no overlap of the 95% confidence intervals) around 0.5 octave above SOAE frequency, the region of primary side-lobes. The PTCs of the three groups did not differ significantly from each other (Fig. 4.2B). In general, STCs appeared to be more sharply tuned than PTCs (Fig. 4.2C).

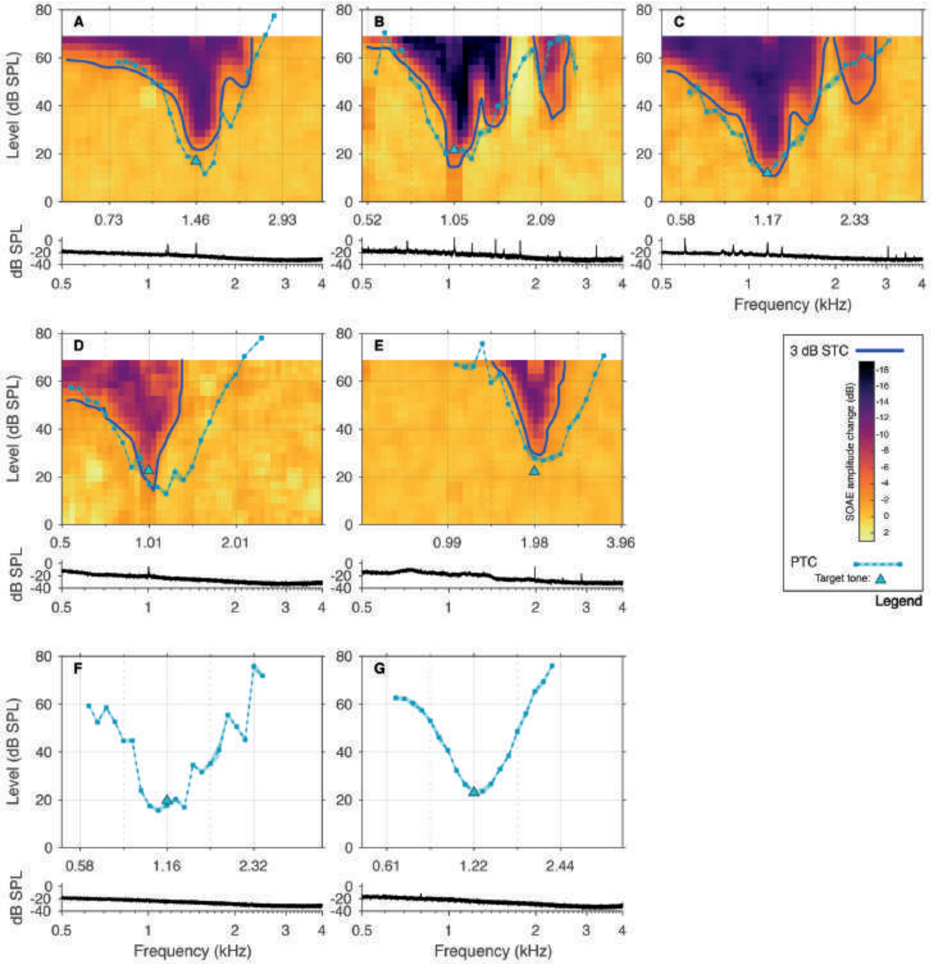


Figure 4.1. Representative STCs and PTCs of seven individuals. In each panel, the horizontal axis displays frequency in kHz. For the plots showing STCs and PTCs, the labels were placed at the emission frequency of interest, and 1 octave below and above that frequency. Each panel on top shows the measurements of frequency selectivity, where (1) the heat map and dark blue contour line indicate 3 dB suppression of the SOAE, and (2) the turquoise dotted line with markers represents the PTC of the same participant. The upper panel shows a grid, labeling ± 1 octave from the SOAE frequency. The frequency and level of the target tone in the PTC measurement are indicated by the green triangle. The corresponding SOAE spectrum is shown underneath the tuning curves. STCs with side-lobes (SL) are shown in panels (A), (B), and (C). Panels (D) and (E) show STC without side-lobes (NSL) while panels (F) and (G) show PTC of participants without SOAEs (No SOAE).

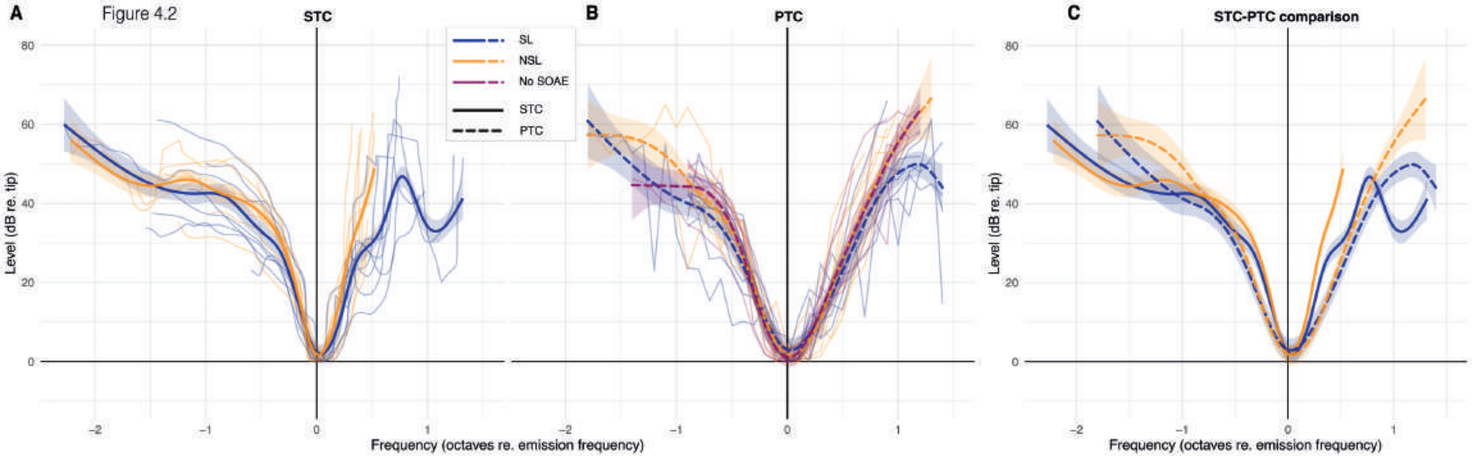


Figure 4.2. Comparison of STCs and PTCs across participant groups. All STCs (panel A) and PTCs (panel B) where curve tips are horizontally aligned to the SOAE frequency and vertically to the tuning curve (TC) tip. Per group, the average tuning curves and confidence intervals (95%) are based on generalized additive model (GAM) likelihood-based fits. TCs of the three participant groups are indicated by different line types and colors: STCs are indicated with solid lines, PTCs with dashed lines. Participants with STC side-lobes (SL) are presented in dark blue lines, participants without STC side-lobes (NSL) are shown in lighter orange lines, and participants without SOAEs (No SOAE) are illustrated with a purple line. Dots underneath both panels indicate the total SOAE count on the STC (panel A) and PTC (panel B) measurement day. Panel C indicates the STC-PTC comparison of the SL and the NSL group. In the SL group, around 0.5 and 1 octave. Above the tip dips can be seen.

Table 4.1. Comparison of STC and PTC measurements across groups. Shown are the median quality factors and slopes of the tuning curves, per group (Fig. 4.3).

Group measurement	Median Q_{10dB}	Median slope low (dB/oct.)	Median slope high (dB/oct.)
SL			
STC	4.25 (sd: 1.39)	-72.42 (sd: 64.75)	73.87 (sd: 16.50)
PTC	3.48 (sd: 1.18)	-49.99 (sd: 19.41)	46.89 (sd: 11.92)
NSL			
STC	5.51 (sd: 1.43)	-88.93 (sd: 38.96)	91.34 (sd: 15.15)
PTC	2.72 (sd: 0.7)	-59.28 (sd: 22.70)	53.84 (sd: 137.35)
No SOAE			
PTC	2.99 (sd: 0.4)	-66.86 (sd: 20.62)	49.48 (sd: 9.27)

STC-tip levels did not differ significantly (*H*(1) = 0.59: p = 0.44, η_H^2 = 0) between the SL (17.22 dB SPL) and the NSL group (14.37 dB SPL). The median PTC-tip levels were also similar between all three groups (SL group: 17.13 dB SPL; NSL group: 16.33 dB SPL, and No SOAE group: 15.58 dB SPL; *H*(2) 1⁄4 0.81, p = 0.67, η_H^2 = 0). Thus, TC-tip levels were similar between STC and PTC measurements and between participant groups.

For all tuning curves, the quality factor Q_{10dB} was evaluated as a measure of the frequency selectivity (Fig. 4.3A, Table 4.1). STCs of participants with SL were similarly sharply tuned compared to the NSL group (median Q_{10dB}: 4.25 and 5.51, respectively; *H*(1) = 1.81, p = 0.18, η_H^2 = 0.05). Also, the frequency selectivity of the PTCs was similar in all three groups. PTCs of the SL group (median Q_{10dB}: 3.48) were similar compared to the NSL group (median Q_{10dB}: 2.72; *H*(1) = 1.12, p = 0.29, η_H^2 = 0.01) and the No SOAE group with participants without SOAEs (median Q_{10dB}: 2.99; *H*(1) = 1.56, p = 0.21, η_H^2 = 0.04). Therefore, the NSL group and the No SOAE group had also no significant differences in PTC sharpness (*H*(1) = 0.33, p = 0.57, η_H^2 = 0).

Second, to answer the question to what extent cochlear tuning may differ from psychoacoustic tuning, we compared Q_{10dB} within each group and between the two methods (STC-PTC comparison; Fig. 4.3A). In the SL group, the Q_{10dB} of STCs and PTCs were not significantly different from each other (*H*(1) = 2.48, p = 0.12, η_H^2 = 0.25). However, in this same group, STCs had significantly steeper low (*H*(1) = 5.34, p = 0.021, η_H^2 = 0.31) and high (*H*(1) = 4.41, p = 0.036, η_H^2 = 0.24) flanks, compared to the low and high flanks of the PTCs (Fig. 4.3B).

The NSL group had significantly sharper STCs than PTCs (*H*(1) = 10.39, p = 0.001, η_H^2 = 0.11). Also, for the NSL group, low and high flanks of the STC were significantly steeper compared to the PTC flanks (*H*(1) = 5.9, p = 0.015, η_H^2 = 0.31 and H(1) = 5.07, p = 0.024, η_H^2 = 0.24, respectively). Details can be seen in Table 4.1.

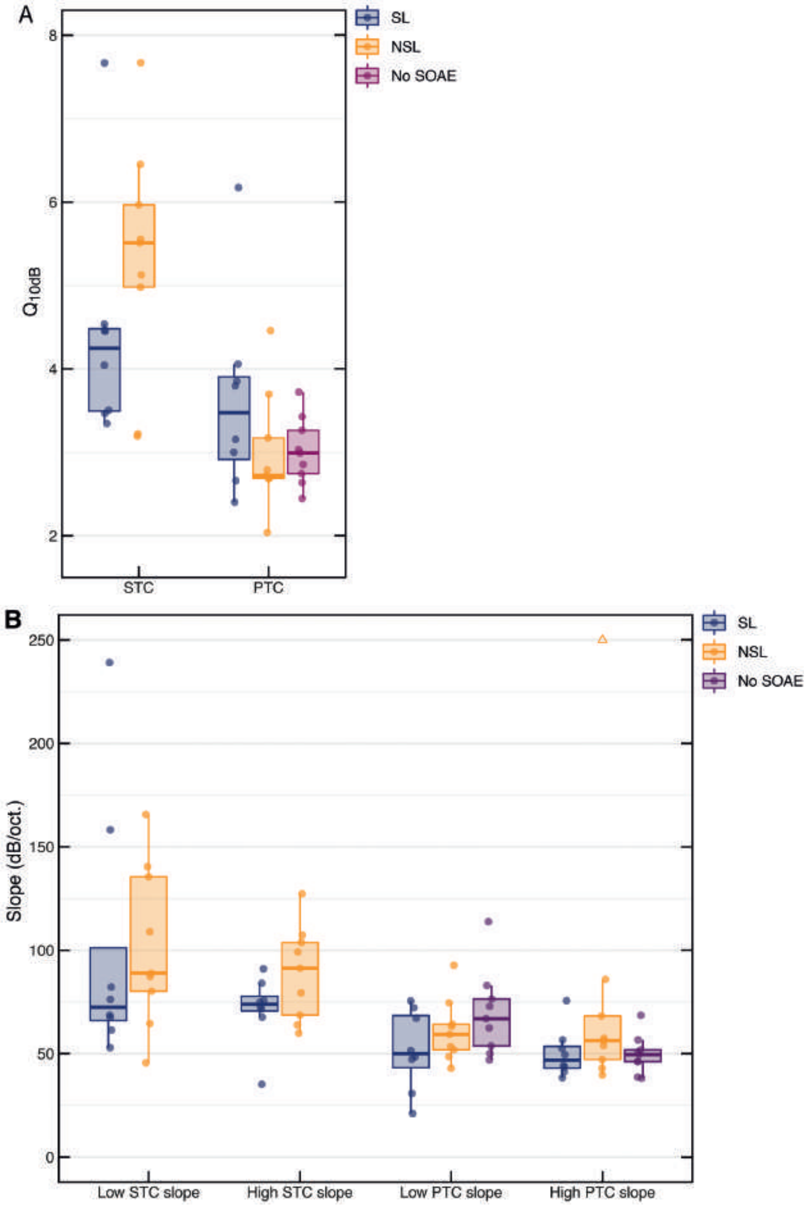


Figure 4.3. Frequency selectivity of the same individuals evaluated via STCs and PTCs. Different subject groups are indicated with different colors: side-lobe group (SL, blue), group without STC side-lobes (NSL, orange), and participants without SOAEs (No SOAE, purple). The middle line in the boxplots indicate the median. The top and lower hinges indicate the first and third quartile. The whiskers extend to the furthest data point within 1.5 times the interquartile range from the hinges. Individual data points are overlaid on the boxplot. In panel A, for each subject, the filter quality Q_{10dB} of the STC is shown as a function of the Q_{10dB} of the PTC. Panel B illustrates the slopes of the tuning curve flanks, below and above the tip (low and high flank, respectively). The orange triangle represents an outlier (at 466.35 dB/octave) for which the axis is not continuously plotted.

4.4 Discussion

While PTCs indicate the limits of the auditory system in resolving two different stimuli (Moore, 1995), probing SOAE-STCs measures the suppressive effect of external tones on a spontaneous emission. STCs require overlap, on the BM, of the vibration patterns corresponding to the SOAE and to the suppressor. Similarly, PTCs are based on interactions between the probe tone and the masker, which also includes interaction on the BM, but may be additionally influenced by the interactions in the central auditory system.

We investigated whether side-lobes that can be found in SOAE-STCs are also present in PTCs. Both STCs and PTCs showed deviation from the standard V-shape. We will refer to side-lobes in the case of STC, and irregularities for PTCs. This terminology seems appropriate as in STC the side-lobes are often very pronounced, whereas for PTC the irregularities may be less distinct. Some STCs show side-lobes, which are additional suppression dips. Primary STC side-lobes appear approximately at 0.5 octave above the SOAE frequency, while secondary side-lobes can be found at about +1 octave (Manley and van Dijk, 2016). In the current study, the comparability between objective and behavioral measures of frequency selectivity was evaluated, with a focus on these irregularities (side-lobes) observed in STCs. We expected to observe PTC irregularities corresponding to the STC side-lobes. To test this hypothesis, we measured PTCs with probe tones at SOAE frequency.

4.4.1 Possible mechanisms of side-lobe generation

SOAEs are believed to correspond to standing waves in the cochlea (Kemp, 1980; Shera, 2003). When forward and backward traveling waves meet in phase, they reinforce each other and form a standing wave. Presumably, a standing wave is present between the base of the cochlea and the characteristic place of an SOAE frequency. SOAEs are formed when frequency-specific, round trip travel requirements within the cochlea are fulfilled. A model of SOAE generation shows that the standing waves corresponding to an SOAE have antinodes at cochlear locations with characteristic frequencies at 0.5 and 1.0 octave above the SOAE frequency (Epp et al., 2015). Interestingly, the side-lobes in STCs are also about 0.5 and 1.0 above the SOAE frequency. This suggests that SOAE-STC side-lobes may be the result of the interaction between the suppressor tone with the standing wave antinode (Manley and van Dijk, 2016).

When the probe tone in the PTC measurements is placed at an SOAE frequency, it presumably generates a standing wave in the cochlea, since the frequency meets the round-trip travel requirements that have to be fulfilled to form the SOAE. We expected that a masker in the PTC measurements would also perturb this standing wave, and thereby, result in side-lobes in the PTC. Although PTC irregularities corresponded to STC side-lobes in some cases (e.g., Fig. 4.1A), this was not always the case (e.g., Fig. 4.1C).

Moreover, side-lobe like irregularities in the PTCs were also present in STCs without side-lobes (e.g., Fig. 4.1D) and even in the absence of detectable SOAEs (Fig. 4.1F). Thus, all three groups showed irregularities in the shape of their PTCs, and PTC irregularities were not systematically related to STC side-lobes. In other words, the relation between SOAE suppression side-lobes and PTC irregularities is not straightforward, suggesting that they may not be generated by the same mechanism.

Some STC measurements indicate that side-lobes may occur in the vicinity of other SOAEs (Fig. 4.1A). Possibly, additional sensitivity dips may be caused by interactions between the probe tone and another SOAE. Moreover, SOAEs can internally interact with each other which can be very diverse and complex. For example, Murphy et al. (1995) describe the suppressive effect one SOAE can have on another one. If the suppressive SOAE is itself suppressed by an external tone, the other SOAE is released from internal suppression. Additionally, when two SOAEs interact (primary SOAEs), distortion product SOAEs (DP-SOAEs) can be generated. Suppressing such a primary SOAE, consequently affects the DP-SOAE, leading to a reduced amplitude (e.g., Burns et al., 1984; Jones et al., 1986; Norrix and Glattke, 1996). In other words, when suppressing a neighboring primary SOAE, the targeted DP-SOAE shows an amplitude decrease as well. Hypothetically, such internal interactions may appear as suppression side-lobes in STCs.

The presence of SOAEs does not only affect the shape of STCs. In fact, previous research has shown that SOAEs can influence the shape of PTCs (Bright, 1985; Micheyl and Collet, 1994; Baiduc et al., 2014). It was suggested that the presence of the emission might be responsible for additional suppression dips (Bright, 1985; Baiduc et al., 2014). Still, it was also shown that in two ears without SOAEs, the PTC deviated from the typical V-shape (Bright, 1985). Indeed, in the current study in ears without any SOAEs, 55.6% of the PTCs demonstrated irregularities (example in Fig. 4.1F). It could be argued that any statement about the presence of SOAEs only concerns observable SOAEs. SOAEs could be present but remain undetected in the noise floor and may still influence PTC recordings. It is very unlikely, however, that the chosen target tone or the noise masker used in a PTC measurement would have interacted with such an undetected emission frequency in several participants, coincidentally.

4.4.2 Potential effects of the PTC masker

Characteristic frequency selectivity measure, as evaluated by PTCs, is influenced by the characteristics of the applied masker. Forward masking paradigms, for instance, reveal sharper tuning compared to simultaneous masking (e.g., Moore, 1978), as the suppressor is not present at the same time as the stimuli. Thus, the temporal presentation of the masker relative to the target stimulus influences the frequency selectivity. Furthermore, whether the masker is a pure tone or a noise band also determines the psychoacoustic tuning results. Tonal stimuli may interact with SOAEs and cause a beating percept

(Wilson, 1980; Wilson and Sutton, 1981; Zurek, 1981). Such cues can be perceived by the participant and therefore, affect the shape of the PTC, potentially leading to the impression of enhanced frequency selectivity. The PTC shape has been reported to be affected by placement of the tonal masker or target tone relative to the SOAE. A target tone at the SOAE frequency can sup- press or synchronize the emission (e.g., Wilson and Sutton, 1981; Zwicker and Schloth, 1984; Long and Tubis, 1988b) and will therefore not lead to frequency beating. Previously, it was suggested that PTCs derived with pure tone masking are more sharply tuned in emitting ears at SOAE frequency than off the SOAE frequency (Bright, 1985). In their study, PTCs were tested with a target tone that matched the SOAE frequency and at least 1 kHz above the emission frequency.

Such kinds of interactions are unlikely to happen when presenting noise bands as maskers, as used in the current study. Because PTCs were obtained with a masker that is spectrally broader than the suppressing tone used in STCs, it was expected that side-lobes observed in PTCs would be less pronounced than for STCs. Thus, the broadness of the noise masker itself may result in broader and/or smoother PTCs, which may obscure the sharpness of PTCs, or the presence of irregularities, in participants with SOAEs. Nevertheless, PTC irregularities were observed, indicating that the width of the chosen masker was not too broad to smooth out dips in the TC.

Micheyl and Collet (1994) compared PTCs at fixed frequencies in emitting versus non-emitting ears. In SOAE emitting ears, PTCs with tonal maskers were significantly sharper at 2 kHz, but not at 1 and 4 kHz. Moreover, ears with weak evoked OAEs showed sharper tuning than ears with strong evoked OAEs. In a more recent paper, it was shown that when using a 3 kHz target tone, PTCs are not significantly influenced by the presence of SOAEs in the investigated ear (McFadden et al., 2018). In that study, neither psychoacoustic tests with notched noise, nor pure tone maskers, and neither simultaneous nor forward masking paradigms were significantly influenced by the presence of SOAEs. Even though McFadden et al. (2018) reported SOAEs over a wide frequency range (0.55 to 9 kHz), it is unlikely that the majority was close to the 3 kHz target tone, since SOAEs are mostly recorded between 1 and 2 kHz (Zurek, 1981; Probst et al., 1991). Thus, it is unlikely that the SOAE and the presented target tone at 3 kHz were interacting and thereby generating any perceivable cue that may enhance the performance of the participant. In the current study, there were also no systematic differences between PTCs in ears with and without SOAEs. Altogether, psychoacoustic tuning does not seem to differ between ears with or without SOAEs.

4.4.3 Comparability of SOAE-STCs and PTCs – A matter of method?

No difference in frequency selectivity (expressed by Q_{10dB}) was observed between the three participant groups, neither for the STC nor for the PTC measurements (Fig. 4.2). In fact, participants without recordable SOAEs showed PTCs that are as sharp as those of participants of the SL and NSL group (Fig. 4.3A). Comparison of the frequency selectivity within subject groups, however, showed sharper STCs than PTCs in the NSL group, whereas Q_{10dB} values for the two measures were similar in the SL group.

In addition to the differences in STC and PTC measurements mentioned above, both measurements are highly dependent on the chosen thresholds for the tuning curve. Here, we defined the STC contour to be an SOAE suppression of 3 dB. For PTCs, we chose the threshold as defined by 70.7% correct target tone detection by the participant (as a result of the 3-AFC procedure). These thresholds are arbitrary and different choices would consequently influence the TC sharpness. We noticed that STCs are sharper tuned when the SOAE-suppression criteria were set to higher levels (comparing 2, 3, and 4 dB of SOAE suppression; data not shown). In fact, studies with evoked emissions showed the same effect. For stimulus frequency OAEs (SFOAEs) and transient evoked OAEs (TEOAEs), STCs were sharper tuned at higher suppression levels (Kemp and Chum, 1980; Zettner and Folsom, 2003; Charaziak et al., 2013). When choosing a higher suppression threshold, towards full SOAE suppression, the STCs may become even sharper tuned then PTCs, increasing the discrepancy between both tuning measures. Thus, there is no reason to expect that frequency tuning of STCs and PTCs would lead to identical results. Note that our experiment was not intended to show that Q_{10dB} quality factors of PTC and STC are identical. Rather, it was set up to investigate the relationship between side-lobes in STCs and irregularities in PTCs (Fig. 4.1).

It is possible that PTC irregularities could result from momentary lapses in attention in the task. We tried to minimize this by presenting a target tone reminder and averaging the threshold over the last six reversals. Evaluating the individual PTCs, the irregularities and dips appear to be present over several neighboring masker frequencies, making it unlikely that we falsely identified, for instance, sudden attention drops as side-lobes. Moreover, we evaluated the reliability of each threshold defining the PTC by calculating the standard deviation across the reversals used to calculate the threshold. While a few irregularities also show enlarged standard deviation at or around the irregularity, most of them did not seem particularly unreliable. It thus seems unlikely that the observed PTC irregularities would be due to momentary attentional deficits.

PTC irregularities were observed, but they did not directly coincide with the presence of STC side-lobes. Moreover, PTC irregularities were also observed in ears without SOAEs present. This suggests that PTC irregularities are not always related to SOAE side-lobes. SOAEs and their STCs are probably determined by the mechanics of the BM and the outer hair cells, whereas the PTC involves the auditory system up to the cortex. Consequently, the origin of side-lobes in STCs and the irregularities in the PTC remains uncertain.

For participants without detectable SOAEs, the probe tone frequency was set to match the probe tone frequency of another subject with SOAE(s). For future studies, the probe tone could be matched to the minima of the audiometric fine-structure. Audiometric measurements with high frequency resolution show ripples that are created by auditory threshold minima and maxima (Elliott, 1958). Such ripples are known as audiometric fine-structure (Thomas, 1975). SOAEs are known to correspond to audiometric such minima (Long and Tubis, 1988b; Mauermann et al., 2004; Heise et al., 2008 and Heise et al., 2009b). Such a minima presumably corresponds to a frequency at which a standing wave could occur, even though there is no detectable SOAE generated.

4.5 Conclusions

Additional PTC dips are not clearly related to STC side-lobes and can also not exclusively be caused by the presence of other SOAEs. In fact, PTCs of participants without SOAEs were just as irregularly shaped. It may be that different mechanisms lead to STC side-lobes and PTC irregularities. Possibly, applying tonal maskers in PTC measurements instead would compare most closely to the suppressive stimulus in the STC measurement. Future research could use tonal maskers to measure PTC that involve the suppression of SOAEs at the peripheral level. Across participant groups, Q_{10dB} of the PTC was similar, independently from SOAE presence or absence.

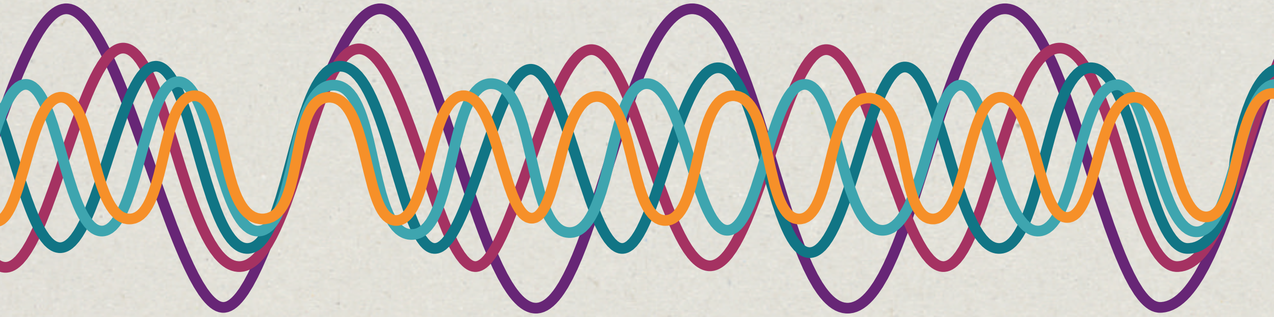
Acknowledgements

This work was supported by the European Glaucoma Research Training Program, Marie Sklodowska-Curie-COFUND, Grant No. 661883 and the Heinsius Houbolt Foundation. This work was conducted in the framework of the LabEx CeLyA (“Centre Lyonnais d’Acoustique,” ANR-10-LABX-0060/ANR-11-IDEX-0007) operated by the French National Research Agency.

---

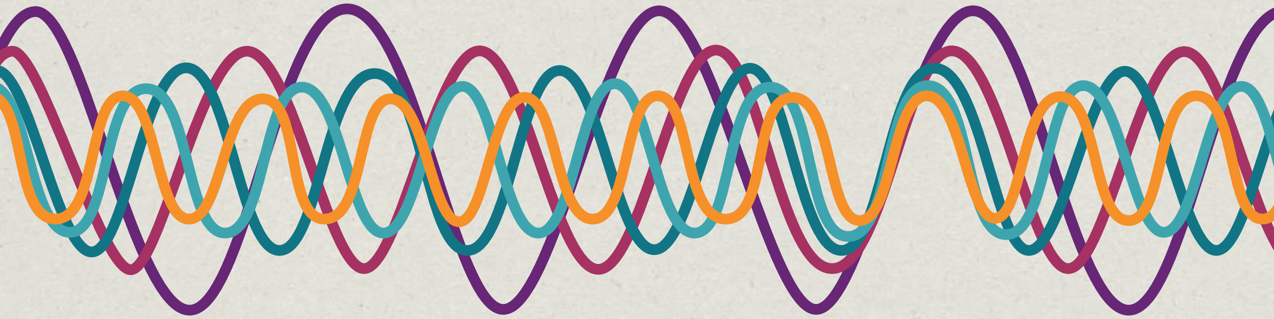
[1] Participants provided informed consent and received monetary compensation for their participation. Both experiments were approved by the Central Ethics Review Committee of the University Medical Center Groningen.

[2] See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0009278 for characteristics of participants that performed the psychoacoustic tuning curve (PTC) measurement.



Chapter 5

Interactions of spontaneous otoacoustic emissions and their influence on auditory perception



Abstract

Although the majority of normal hearing participants have spontaneous otoacoustic emissions (SOAEs), which may significantly affect psychoacoustic measurements such as auditory threshold and frequency selectivity, SOAE-screening is not routinely done prior to such assessments. This review outlines the extensive complexity of interactions of SOAEs (1) with each other and (2) with external tones. It is described how these interactions may cause inter-participant variability in psychoacoustic measurements and how these interactions can influence participants' perception during psychoacoustic tasks.

5.1 Introduction

One fascinating phenomenon of the human ear is the ability to produce sound. These sounds, or otoacoustic emissions (OAEs) can be recorded in a person's ear canal. When emissions are recorded in absence of external acoustic stimulation, they are termed spontaneous OAEs (SOAEs). It has been shown that people with SOAEs have lower hearing thresholds across frequencies compared to those without SOAEs (McFadden and Mishra, 1993). The frequency of SOAEs in humans coincides with that of local minima in auditory thresholds (e.g.: Smurzynski and Probst, 1999). Moreover, psychoacoustic measures of frequency selectivity were shown to differ significantly between ears with and without SOAEs. Some studies have shown that participants with SOAEs appear to have sharper frequency selectivity when probed psychoacoustically (e.g.: Micheyl and Collet, 1994). Moreover, within participants the selectivity appears to be sharper at an SOAE frequency compared to frequency regions without SOAEs (Bright, 1985). However, recent studies indicate that correlations between SOAEs and psychoacoustic tests may also be less straightforward (McFadden et al., 2018; Engler et al., 2022).

When presenting a tone close to the emission frequency, as it could happen during audiometry and measurements of frequency selectivity, the SOAE and the presented tone may interact. These interactions can lead to emission amplitude and frequency changes, where the latter may be perceived by the participant as *beating* or *roughness* (Wilson, 1980). In addition to interactions between an external stimulus and an SOAE, multiple SOAEs within an individual ear may also interact. Such interactions reflect a change of acoustic vibration within the inner ear. This raises intriguing questions: Do participants perceive such interactions in psychoacoustic measurements? And to what extent will SOAE interactions be reflected in the outcome of psychoacoustic experiments?

In this review these questions will be addressed. First, the interactions of SOAEs with each other (internal, section 2.1) and with presented stimuli (external, section 2.2) are considered. Then, the effects of these interactions on behavioral auditory measurements are further discussed (section 3). This review will facilitate our understanding of inter-participant variability and the comparability between measurements in general. The awareness of SOAE interactions is of importance especially when performing psychoacoustic measurements with tones close to perception threshold. The recognition of SOAEs and their interactions is important to choose suitable measurement parameters (e.g.: a probe tone frequency) and for the correct data interpretation, especially when deviations (at intra- and inter-participant level) are present.

SOAE characteristics, prevalence, and generation

SOAEs can be understood as the ear mechanics in reverse: the signal is generated in the inner ear and is transmitted through the middle ear to the ear canal, where it can be recorded with a sensitive microphone as an otoacoustic emission. In a time-averaged spectrum, SOAEs appear as peaks above the microphone noise floor. Here, the sound component that gives rise to a single peak is referred to one SOAE. In the spectrum, recordable SOAEs usually appear as narrow peaks with only small and random fluctuations in amplitude and frequency (van Dijk and Wit, 1990b). Thus, the SOAE itself has pure tone characteristics (Schloth and Zwicker, 1983; Bialek and Wit, 1984). Multiple SOAEs may be recorded from a single ear.

The presence of SOAEs represents a common phenomenon in normal-hearing human ears. Between 72 and 90% of human ears emit them (Talmadge, et al., 1993; Penner and Zhang, 1997; Pasanen and McFadden, 2000). The occurrence of SOAE peaks is individual-specific and very stable over years (Burns 2009, 2017). In females, SOAE prevalence is higher and more SOAEs per ear are emitted compared to males (e.g.: Strickland et al., 1985; Bilger et al., 1990; Moulin et al., 1993; Whitehead et al., 1993; Penner and Zhang, 1997). In fact, it was shown that the incidence in female ears to emit SOAEs is between 21 and 36% higher compared to males (Moulin et al., 1993; Penner and Zhang, 1997) and that approximately twice as many spectral SOAE peaks are found in ears of females compared to males (Talmadge et al., 1993). Thus, it is very likely that a normal hearing participant in an auditory experiment emits SOAEs, especially when being female. Interestingly, hormone-status appears to affect SOAE occurrence. When estrogen and progesterone levels are low, emissions of females are more male-like (McFadden et al., 2018).

Most of the SOAEs are found at frequencies below 4 kHz (Moulin et al., 1992) with the majority being recorded within the frequency range of 1 to 2 kHz (Zurek, 1981; Probst et al., 1991; Talmadge et al., 1993). Presumably, the reverse transfer function of the middle ear favors the transmission of this frequency range (Probst et al., 1991). SOAE prevalence (Whitehead et al., 1993) and frequencies (Chan and McPherson, 2001; Engler et al., 2020a) were shown to be different between language groups which could be caused by differences in ear canal and middle ear properties that are related to ethnicity (Shahnaz and Davies, 2006; Shahanz, 2008).

SOAEs are considered the most direct evidence of active sound amplification in the ear. Early models described SOAEs as signals arising from self-sustained oscillators (e.g.: Bialek and Wit, 1984). These models can explain the sinusoidal character of the SOAE signal, as well as their random amplitude and frequency fluctuations, when internal noise is incorporated in the oscillator. Moreover, the models capture the influence of external tones on the SOAE properties, such as phase-locking and suppression (Talmadge et al., 1990; Long and Tubis, 1988b; van Dijk and Wit, 1988). To some extent, they can also describe interactions between multiple SOAEs in an individual ear (Murphy et al., 1995). The rationale behind the self-sustained oscillator is that hair cells amplify acoustic vibration on the basilar membrane (BM). This active feedback may cause the BM to become unstable for certain frequencies, which would result in the spontaneous generation of sound within the inner ear, namely SOAEs, as initially predicted by Thomas Gold (Gold, 1948).

More comprehensive cochlear models, that describe forward and backward traveling waves on the BM, have explained SOAEs to result from standing waves (Kemp, 1980; Shera, 2003; Epp et al., 2015). Here, backward waves are reflected when they meet the transition from the inner to the middle ear, that is, the oval window. Forward waves are scattered back near the characteristic frequency place on the BM. A standing wave may form when the round-trip travel time from the oval window to the characteristic place and back matches an integer multiple of the oscillation period. This condition is met for specific frequencies only. Hence, these global models can account for the characteristic spacing of SOAE peaks in a spectrum, and for the modulating effect of middle ear properties on SOAE amplitude and frequency (Shera, 2003; 2022).

5.2 SOAE interactions

SOAE spectra are typically stable over years (Burns, 2017). Nevertheless, SOAEs dynamically interact with each other and with external stimuli, causing changes in SOAE frequency and amplitude. Internal interactions (between individual emissions) and external interactions (between the emission and an external auditory stimulus) will be discussed in this section.

5.2.1 Internal interactions

Within the same ear multiple SOAEs can interact with each other. The interacting relationship between SOAEs is not always easy to detect when simply monitoring them, because the interaction is relatively stable in unmanipulated conditions. In some cases, these internal interactions can be identified, for example when one *interaction partner* is manipulated by an external tone that perturbates the stable SOAE-SOAE interaction (further discussed in section 2.2).

Generation of additional emission peaks

Two SOAEs can interact and produce distortion products that are recorded as additional smaller SOAEs (Burns et al., 1984; Jones et al., 1986; Frick and Matthies, 1988). Here, they are termed distortion-product SOAEs (DP-SOAEs), analogous to distortion product otoacoustic emissions (DPOAEs) which are produced by two external tones. DP-SOAEs are distortion products that result from the interaction between two SOAEs or between an SOAE and an external tone. DP-SOAEs were identified in 12.5% of participants with multiple SOAEs in one ear (Frick and Matthies, 1988). Here, the focus is on DP-SOAEs that are a product of SOAE-SOAE interaction only.

DP-SOAEs follow the rules of DPOAE generation and consequently a specific relation exists between the frequencies of the two-coexisting *primary* SOAEs (SOAE_P) and the produced DP-SOAE. A spectral analysis of the SOAE recording is essential to identify

DP-SOAEs, when evaluating the frequency spacing between multiple emission peaks. When three SOAEs of equal frequency distance between each other are recorded, it is very likely that two are $SOAE_P$ and one is the DP-SOAE (van Dijk and Wit, 1998a). DP-SOAEs are mainly generated at cubic frequency relations, such as $2f_1$-f_2 or $2f_2$-f_1 (van Dijk and Wit, 1998a). The strongest DP-SOAEs are generated at $2f_1$-f_2 (Talmadge et al., 1983; Burns et al., 1984).

Since DP-SOAE levels are lower than $SOAE_P$ levels (Zwicker, 1981; Burns et al., 1984), the $SOAE_P$ need to be of sufficient amplitude in order to generate recordable DP-SOAEs that exceed the noise floor. On average, the level of the DP-SOAE is 14 dB below the level of the f_1 $SOAE_P$ and 10 dB below the level of f_2 $SOAE_P$ (Burns et al., 1984). Therefore, *the smallest* of the emission triplet in a spectrum is most likely to be a DP-SOAE. The level of a DPOAE generated by two external tones is at least 40 dB below the level of the external tones, when f_1 and f_2 are of equal level (Norrix and Glattke, 1996). Within the same frequency range, relative to the levels of the primaries, DP-SOAEs are (at least) 21.5-30 dB lower, compared to DPOAEs produced by external tones (cf. Burns et al., 1984). DP-SOAEs can also be identified by their peak width. The small frequency fluctuations of the two $SOAE_P$ will lead to a broader peak of the DP-SOAE (van Dijk and Wit, 1998a). Thus, the *widest emission peak* in such an emission triplet probably corresponds to the DP-SOAE.

If an SOAE happens to be close to the frequency of the cubic distortion product of two other SOAEs, it is possible that an *independent* SOAE synchronizes to a DP-SOAE (van Dijk and Wit, 1998b). In such cases the synchronized SOAE is under constant influence of the DP-SOAE. Thus, changes in $SOAE_P$ (that produces the DP-SOAE) will cause changes in the DP-SOAE and the synchronized SOAE. The SOAE may or may not be synchronized any longer to the DP-SOAE.

Besides distortion products, $SOAE_P$ can also produce harmonics that are recorded as multiple SOAEs which all are of nearly the same amplitude (Talmadge et al., 1993). They can be identified by their stable doubling in frequency in relation to the $SOAE_P$.

Amplitude and frequency shifts of emission peaks

Besides the generation of distortion products, neighboring SOAEs can also influence each other's amplitude and frequency. This type of internal interaction is also stable during unmanipulated conditions, thus these interrelations are not easy to identify when simply monitoring SOAEs.

SOAEs can, for instance, suppressively interact with each other (Burns et al., 1984; Murphy et al., 1995). During this SOAE-SOAE interaction one emission can also be affected in amplitude *and* frequency by the other SOAE. In general, the stronger SOAE (1) suppresses the weaker emission, and (2) pulls the weaker SOAE closer to itself. When the internal suppressor is now itself suppressed by an external tone, the internal suppression will be canceled (suppression *release)*. Consequently, the previously suppressed SOAE will increase in amplitude and may shift in frequency. The phenomenon of *released internal suppression* will be further discussed in the following paragraph of external interaction.

5.2.2 External interactions

As already mentioned, an external tone can affect the amplitude and the frequency of SOAEs. In the following, it will be explained how such external interactions between SOAEs and presented stimuli occur and how these interactions can be identified.

Generation of additional peaks

When an SOAE and a pure tone interact, distortion products can be generated (Frick and Matthies, 1988; Norrix and Glattke, 1996). The principle is similar to that of earlier discussed DP-SOAEs, only that in this case not *two* SOAE_P, but one SOAE_P and one external tone produce the DP-SOAE. DP-SOAEs can only be generated when the external tone is of sufficient level to interact with the SOAE_P (Frick and Matthies, 1988). Interestingly, up to a certain extent the SOAE_P needs to be suppressed in order to generate the maximal DP-SOAE amplitude (Norrix and Glattke, 1996; see next paragraph).

Amplitude suppression of emission peaks by presented tones

As mentioned before, during the interaction between an external pure tone and the SOAE, emissions can be suppressed (e.g.: Burns et al., 1984; Rabinowitz and Widin, 1984; Jones et al., 1986). Tones of low levels can already suppress the emission when their frequency is close to the SOAE frequency, the suppressive effect increases with increasing tone level. This suppressive effect of external tones on the SOAE amplitude is deliberately used in objective measurements of cochlear frequency selectivity derived by suppression tuning curves (STCs; Schloth and Zwicker, 1983; Engler et al., 2020a).

Suppressing an SOAE with a pure tone can reveal internal suppressive dependencies between two interacting SOAEs, especially when the suppressing SOAE is becoming suppressed by the tone. While this suppressing SOAE is now itself suppressed, the other SOAE increases in amplitude. Once again, this phenomenon of *released* (internal) SOAE suppression (Burns et al., 1984; Murphy et al., 1995) may, besides changes in amplitude, also cause shifts in frequency of a previously suppressed SOAE.

SOAEs recover from suppression within milliseconds after the external tone is turned off. The effect of *internal suppression* can also be visualized by the amplitude recovery of two interacting SOAEs that were both suppressed by an external stimulus. Interestingly, after the external suppressor is turned off, one SOAE may show a fast amplitude recovery with an initial overshoot in amplitude-rise, while the other recovers relatively slowly (Murphy et al., 1995). The slower recovering emission is that one that functions as an internal suppressor. Once this emission is fully recovered, it suppresses the other SOAE again.

A release of SOAE suppression can also be observed when presenting two external tones. When both external tones are of higher frequencies than the SOAE, the tone closer in frequency to the SOAE can suppress the emission. This suppression effect of the nearest tone may be diminished, however, when the additional second tone is presented (Rabinowitz and Widin, 1984).

The direct suppression of a DP-SOAE affects the $SOAE_P$ (e.g.: Burns et al., 1984). Furthermore, DP-SOAEs can also be affected indirectly by external tones. An external tone that directly suppresses the $SOAE_P$ can also indirectly suppress the linked DP-SOAE (e.g.: Burns et al., 1984; Jones et al., 1986; Norrix and Glattke, 1996). Firstly, however, when increasing the level of the external tone, the level of the DP-SOAE increases while the $SOAE_P$ level is reduced (Norrix and Glattke, 1996).

In summary, an external tone can interact with an $SOAE_P$ and generate a DP-SOAE, in which the relation between $SOAE_P$ amplitude and DP-SOAE amplitude is highly nonlinear (Norrix and Glattke, 1996). For low-level tones, the DP-SOAE appears and the $SOAE_P$ is marginally suppressed. Maximum DP-SOAE level is related to a significant $SOAE_P$ suppression. When the external tone reaches a critical level, however, the $SOAE_P$ will be suppressed which consequently causes the disappearance of DP-SOAE.

Interestingly, the DP-SOAE attenuation appears to be steeper when the $SOAE_P$ is suppressed, as compared to DPOAEs generated by two external tones (Norrix and Glattke, 1996). In fact, Norrix and Glattke (1996) found these trends in DP-SOAEs so unique that they cannot be compared to DPOAEs, and suggested different mechanisms generating the distortion products.

Frequency shifts of emission peaks caused by the presence of external tones

Besides SOAE amplitude changes, external tones may also cause shifts in SOAE frequency. The frequency shift can be typed as frequency *synchronization* (locking, phase locking) or *unstable shifts* (pulling and pushing). The occurring type depends on the frequency and level of the presented pure tone.

An SOAE may synchronize to an external tone that is close in frequency (Wilson and Sutton, 1981; Zurek, 1981; Bialek and Wit, 1984; Zwicker and Schloth, 1984; Long and Tubis, 1988a). During this interaction, the emission frequency can be shifted over several Hz (Rabinowitz and Widin, 1984). Still, as indicated in Figure 5.1, the frequency range over which tones affect the SOAE is narrow and the level of the external tone is decisive for the appearance of SOAE changes (Zwicker and Schloth, 1984; van Dijk and Wit, 1990b; Long and Tubis, 1988a; Long, 1998). A DPOAE can also synchronize an SOAE. In a study by van Dijk and Wit (1990b), primary tones of only about 30 dB SPL produced a DPOAE that could synchronize a nearby SOAE. Such weak primaries would not produce detectable DPOAEs in absence of the SOAE. Nevertheless, a weak nonlinear distortion product is apparently generated in the inner ear that can interact with a nearby SOAE.

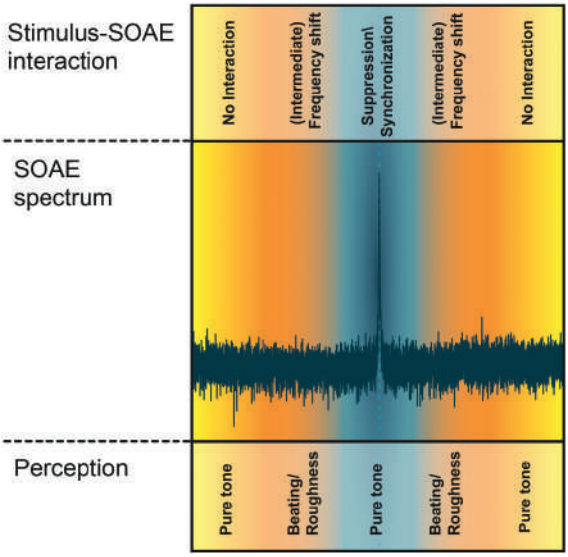


Figure 5.1. Illustration of the interaction between an SOAE and an external tone. The presence of an external tone can influence the level and frequency of an SOAE. The spectrum in the middle shows one SOAE peak, that is well above the noise floor. The SOAE peak-frequency is indicated by a dashed line. The top panel shows the effect the presented tone has on the emission, while the bottom shows how this interaction is perceived by the emitter. Any low-level tone (± 25 dB SPL) presented close to the SOAE (within the blue area), suppresses the SOAE amplitude or locks the emission frequency. This condition is perceived as a pure tone. Tones of the same level, but further off the SOAE frequency can shift the emission frequency, an interaction that may be perceived as beating (orange area). Here, the tone does not completely suppress the SOAE, unless the presented level is increased. When the external stimulus is even further away from the SOAE frequency, the emission is not influenced (yellow area). Consequently, only the external tone is perceived. External tones of higher levels influence the SOAE over a broader frequency range. Thus, the width of the illustrated areas of interaction depends on the level of the external tone.

SOAE frequency synchronization will not occur when the external tone is just below the sufficient level of frequency locking or further off the emission frequency (Wilson and Sutton, 1981; Zurek, 1981; van Dijk and Wit, 1990b). In these cases, the SOAE will in fact *escape* from locking and oscillate at its original frequency again. In the intermediate state, the SOAE oscillates occasionally at its original (unmanipulated) frequency and randomly jumps back to synchronization. This state may be perceived as a rather irregular and rough or *tremble-like beating* or *ringing* (Wilson and Sutton, 1981; Wilson, 1980; Zurek, 1981). In this review such perceived frequency shifts are termed beating.

In this state, multiple or combined tones can be perceived in the emitting ear (Zurek 1981; Long 1998). Presumably, earlier studies before the discovery of SOAEs, describe the perception of the external interaction as *monaural diplacusis* (Flottorp, 1953; Ward, 1955), from now on referred to as diplacusis. With the increase in stimulus amplitude the beating-rate of the diplacusis decreases (Flottorp, 1953; Ward, 1955; Zurek 1981), presumably since the SOAE synchronizes to the stimulus. It was described that the perception of beating depends also on the frequency separation between the presented pure tone and the SOAE. When presenting two tones, the beating is faster at larger frequency separations, which makes its detection more difficult (Moore, 2013). Presumably, an external tone that interacts with an SOAE has similar perceptual effects as the described beating by Moore (2013). Since a nearby tone locks/suppresses the SOAE the intermediate state, between no sufficient effect on the SOAE and just before SOAE locking or suppression, may cause the perception of rather slow beats. In such a state the SOAE may be partly locked to the frequency of the external tone and switch back to its original frequency.

5.3 Psychoacoustic measures affected by the presence of SOAEs

Usually, individuals can not perceive their own SOAE(s). However, when presenting a low-level pure tone that interacts with an SOAE, the participant may perceive a stimulus that deviates from the perception of a simple pure tone (as described in the previous section). Such an interaction can therefore influence the results of psychoacoustic tests and also lead to unexpected variation between participants.

For example, the presence of SOAEs seems to influence pitch perception. In binaural experiments, musicians perceived tones above their SOAE frequency higher in pitch and tones below the emission frequency lower in the emitting ear compared to the ear without SOAEs (Köhler and Fritze, 1994). The precise frequency discrimination of about ±30 Hz around the SOAE frequency is decreased (Köhler and Fritze, 1994). With increasing tone intensity, the SOAE is pulled towards the external tone (Long, 1998), enhancing the effect of changes in tone perception. When the SOAE frequency is locked to the pure tone frequency, the beating percept is relieved to that of a rather continuous tone (Long and Tubis, 1988a). In this section, various measurements will be discussed in which these interactions may play a role.

5.3.1 Audiometric threshold fine structure

Audiometry with high frequency resolution revealed a *ripple in the audiogram*, consisting of threshold minima and maxima (Elliott, 1958). The *ripples* refer to auditory threshold minima and maxima, also called the *threshold fine structure* or *microstructure* (Thomas, 1975). Later, it was shown that SOAE center frequencies correspond to the frequency of the threshold minima (Horst et al., 1983; Long and Tubis, 1988b; Mauermann et al., 2004; Heise et al., 2008, 2009b).

The relation between SOAEs and lower audiometric thresholds (Zwicker and Schloth, 1984; Long and Tubis, 1988b; Horst et al., 1983) suggests that fine structure is indicative of a healthy cochlea (Epp et al., 2010). It has to be noticed that SOAEs of small amplitudes may be present, but remain undetected (Heise et al., 2009). This may explain why threshold fine structure can also be measured apparently *unassociated* to SOAEs, or in absence of any recordable SOAEs (Zwicker and Schloth, 1984; Long and Tubis, 1988a, b; Smurzynski et al., 2001; Baiduc, 2014). Audiometric threshold minima and maxima were described as a result of cochlear standing waves (Kemp, 1978). When forward and backward traveling waves integrate, they can create standing waves. Presumably, threshold fine structure minima correspond to frequencies at which a standing wave may occur. In fact, models indicate that the spacing of threshold minima resembles the spacing between neighboring SOAEs (Kemp, 1979b; Talmadge et al, 1998) supporting the standing wave model for SOAE generation.

The perception of pure tones at low levels near threshold may be influenced by their interaction with SOAEs. When the frequency of a presented pure tone is close to a fine structure minimum, the SOAE synchronizes to the stimulus (Long and Tubis, 1988a; Fig. 5.1). Consequently, the percept is that of a tone (Long and Tubis, 1988a.; Long et al., 1991). At threshold maxima, however, the presented tone is further off SOAE frequency. The tone causes shifts in SOAE frequency but does not synchronize it. Consequently, the simultaneous presence of the stimulus and the SOAE causes a beating percept (Long and Tubis, 1988a). That may also explain why a threshold minimum may be associated with diplacusis (Zurek, 1981; Wilson, 1980), where the presented tone may interact with an SOAE.

Fine structure minima can be as deep as 15 dB (Baiduc et al., 2014). Although the audiometric fine structure has a strong correlation to SOAEs (e.g.: Smurzynski and Probst, 1999), the depth of the threshold dip is not directly related to the SOAE amplitude (Long and Tubis, 1988a; Smurzynski et al., 2001). Still, fine structure and OAEs seem to be linked. For example, normal hearing participants with unusual small amplitudes of transient evoked OAE (TEOAs) also had significantly less-pronounced audiometric fine structure (Kapadia and Lutman, 1999). A fine structure minimum is generally not observed at the frequency of a DP-SOAE. (Burns et al., 1984). This reflects the fact that the frequency of a DP-SOAE is not determined by the presence of a cochlear standing wave. Instead, it is determined by the frequencies of the primary SOAE.

The duration of stimuli influences the depth of the threshold fine structure (Smurzynski and Probst, 1999). Longer tone durations (e.g.: 160-320 ms) led to clear fine structure minima at SOAE frequency. The fine structure minimum becomes wider when tones of 40-80 ms are presented. No fine structure dip is seen, when presenting an even shorter tone (20 ms). It was reported that not only SOAE frequency-beating, but also SOAE-locking favors the participants' perception of the presented tone (Smurzynski and Probst, 1999). Longer tone duration thus led to longer interaction-times, which can be more easily detected and explains why fine structure vanishes with decreasing tone duration.

5.3.2 Psychoacoustic tuning curves

Frequency selectivity can be measured with psychophysical tuning curves (PTCs). Then, the participant's task is to detect a target stimulus (probe tone) while a masker (e.g.: tone or narrowband noise) is present. The individual's frequency selectivity is then visualized as the masker level for which the probe tone is just detectable. As a function of masker center frequency, this forms a roughly V-shaped tuning curve (TC). When masker and probe tone are very similar in frequency (at the tip of the V-shape), the masking effect is greatest - a phenomenon that is known for a long time already (Wegel and Lane, 1924).

When measuring PTCs with a probe tone at SOAE frequency, the PTC-tip (most sensitive threshold) is usually just above SOAE frequency (Baiduc et al., 2014), as also shown for STCs (e.g.: Manley and van Dijk, 2016). To which extent the presence of SOAEs may influence the outcome of TC-measurements, remains an active field of research. It is very likely that the frequency of the target tone relative to the SOAE and the masker-choice has an impact on the PTC shape in spontaneous emitting ears.

When the TC is influenced by SOAE interactions, additional tips may lead to a wider curve, especially when multiple SOAEs are present (c.f. Baiduc et al., 2014). On the other hand, some studies have shown that PTCs appear to be generally sharper (greater Q_{10dB}) in ears with SOAEs (Baiduc, 2014) and when being measured close to an SOAE (Bright, 1985; Micheyl and Collet, 1994). Micheyl and Collet (1994) showed that frequency selectivity was not significantly increased at probe tone frequencies of 1 and 4 kHz in ears with SOAEs. However, PTCs around 2 kHz were significantly sharper, which was also the mean SOAE frequency recorded in their study. Thus, it may reasonably be hypothesized that frequency selectivity is enhanced when the probe tone is close to the SOAE frequency. In Chapter 4 of this thesis, however, it is shown that Q_{10dB} values are similar regardless of SOAE presence or absence (Engler et al., 2022).

Moreover, the typical V-shape of the PTC can be rather rippled or W-shaped, possibly due to the presence of SOAEs (Baiduc et al., 2014). When multiple SOAEs (including their distortion products or harmonics) are present and they interact with the presented pure tone, the TC-shape may become rather complex. This may happen for PTC and STC measurements (e.g.: Bright, 1994; Baiduc et al., 2014). In both types of TC measurements an additional SOAE (or its product) needs to fall within the flanks of the TC, in order to affect the curve shape. However, when measuring PTCs and STCs in the same individual, the TC irregularities do not always match (Engler et al., 2022). This suggests that different mechanisms can cause such irregularities in PTC and STC shapes.

5.3.3 Detection of the just noticeable intensity differences between tones

Experiments in which the just noticeable difference (JND) in intensity is determined, also may use tonal stimuli. It was shown that discriminating intensity JNDs does not significantly differ between normal hearing participants with and without SOAEs. For 1-kHz tones at 20, 40, and 60 dB SL, mean JNDs of both groups were similar (Probst and Harris, 1996; Smurzynski and Probst, 1999). However, participants with SOAEs had significantly higher intra-test variabilities for low target tone levels (20 dB SL) and less test variability at higher tone levels (60 dB SL). When presenting tones close to an SOAE frequency (30 Hz below), the variability of the results across participants was increased (Probst and Harris, 1996). During these measurements the SOAE probably interacted with the presented tone. In fact, the interaction in some cases became so severe that the task could not be performed (Smurzynski and Probst, 1999). The confusion of the participants during this task is presumably caused by the perceived interaction of the SOAEs and a low-level external tone. While external tones of higher level rather suppress the SOAE, which does not result in a perceivable interaction as beating. Thus, SOAEs affect an intensity discrimination task at relatively low stimulus levels only.

5.3.4 Detection of tone amplitude modulations

When a pure tone (carrier) shows temporal fluctuations in its envelope, its amplitude is modulated. The lowest sinusoidal amplitude modulation that is perceived, is called the modulation detection threshold (MDT) for a given carrier tone. With the spectrum of an amplitude modulated (AM) sound consisting of a carrier and two side-bands (± the modulation frequency from the carrier) it was shown that the perception of the AM is related to the position of the side-bands relative to the threshold fine structure (Heise et al., 2009b). The MDT was lowest when the carrier was presented at a fine structure maximum, while the side-bands were at fine structure minima. MDTs were higher for the opposite condition, when the side-bands are located at the fine structure maxima.

Potentially, the interaction between an SOAE and an AM-stimulus affects the detectability of the modulation. If the carrier of an AM stimulus is close to the fine structure minima, an SOAE may interact with the stimuli and be synchronized to it (Heise, et al. 2009a, b). Since the SOAE adds in phase to the stimulus, the carrier modulation depth is reduced, (Long, 1993). Hence, the relative strength of the AM modulation will be decreased, making the detection task for the participant more difficult. Hence, SOAE synchronization may explain the observed effect of higher MDTs at threshold minima (Heise, et al. 2009a; b).

5.3.5 Detection of stimulus gaps

The gap detection task is a method to test temporal resolution. In such measurements the participant is asked to indicate whenever the small pause in a stimulus is detected. Gap detection in broadband noise (0.1-12 kHz, 500 ms) is only significantly different between participants with and without SOAEs at low stimulus levels of 10 dB SL, not for higher levels (Smurzynski and Probst, 1999). Participants without SOAEs had lower gap detection thresholds. The presence of multiple and large SOAE peaks seems to obstruct the gap detection at low stimulus levels, resulting in higher thresholds (at 10 dB thresholds for participants with SOAEs: 9.4 ms; non-emitters: 7.7 ms; Smurzynski and Probst, 1999). It was suggested that the gap is partly masked (filled) by the SOAE and consequently is less audible (Smurzynski and Probst, 1999).

When presenting a low-pass noise stimulus (covering 0.1-4 kHz, at 10 dB SL) while partially filling the gap, participants with SOAEs had lower detection thresholds compared to empty gaps (Smurzynski and Probst, 2000). In this study both gap types had the same duration. Empty gaps were defined by the absence of any noise stimulation, partly filled gaps were of lower depth decrement in which the level was depending on the presentation level. The *internal noise* created by the SOAE may be suppressed when the gaps are filled. This is another indication that SOAEs indeed function somewhat like an *internal masker* as initially named by Long and Tubis (1988a). When presenting higher noise levels, the SOAE is probably suppressed which then leads to very similar gap detection thresholds as in people without SOAEs (Smurzynski and Probst, 1999). Note that the gap was centered to the broadband noise midpoint, not to a specific SOAE (Smurzynski and Probst, 1999; 2000). Nevertheless, the detection of a gap was affected by the presence of SOAEs.

In summary, SOAEs appear to partially fill gaps that are inserted in a noise stimulus. As a result, the detection of the gap is made more difficult. For higher stimulus levels, however, this effect is reduced as the SOAE may be suppressed.

5.4 Factors changing SOAE characteristics

Especially in behavioral studies it is often difficult to explain inter- and intra-participant variability in test results, as multiple factors could account for such differences (e.g. training status, fatigue). So far, the focus was on SOAEs that can explain variability in psychoacoustic studies. However, changing SOAEs characteristics other than by acoustic stimulation can additionally influence the variability, which is especially important in group comparisons and long-term monitoring studies. The following paragraphs point out some factors that can influence the characteristics of emissions and thus deserve special attention when conducting psychoacoustic measurements.

5.4.1 Body position

It is known that pressure alterations due to changes in body position affect the conduction of sounds through the middle ear (Macrane, 1972; Wilson and Sutton, 1981; Büki et al., 2002) and thus SOAE characteristics (Bell, 1992; de Kleine et al., 2000). When tilting a participant head down the intracochlear fluid pressure (intracochlear pressure; ICoP) increases, causing an increase in the stiffness of the cochlear windows resulting in SOAE amplitude and frequency changes (de Kleine et al., 2000). It was shown for the gerbil model that ICoP alterations result in changes of the stiffness in the stapes system (Büki et al., 2002). An inverted body position of 30 degrees only, can cause SOAEs to appear and disappear, and can cause SOAE frequency shifts in both directions. For frequencies below 2 kHz, the largest changes were observed (de Kleine et al., 2000). The middle ear impedance can thus change the characteristics of emissions (Avan et al., 2000).

Of course, such changes of SOAE characteristics can affect psychoacoustic measurements. For audiometric tests for example it was shown that auditory thresholds increase (Macrane, 1972) and fine structure minima and maxima interchange (Wilson, 1980). As a consequence of body tilt, SOAEs can also temporarily become audible to the individual itself (Schloth and Zwicker, 1983). (Note, the sudden perception of the SOAE during the trendelenburg position was presumably also described as tinnitus in early studies (e.g.: Wilson, 1980; Wilson and Sutton, 1981). Detecting one's own SOAE may lead to confusion for the participant during acoustical tasks, such as pure tone audiometry or PTC-tasks.

5.4.2 Ototoxic drugs

The consumption of ototoxic drugs can also affect SOAE characteristics and psychoacoustic performance. Aspirin, for instance, can temporarily reduce and even eliminate SOAE amplitudes and lead to reversible frequency shifts (McFadden and Plattsmier, 1984; Long and Tubis, 1988b). Apparently, Aspirin concentration affects the micromechanical parameters, in a way that the active cochlear mechanism (which also generates SOAEs) is reduced (Long and Tubis, 1988b). Consequently, SOAE amplitudes decrease which leads to a relief of some sort of *internal masking* by the emission (Long and Tubis, 1988a). In most cases threshold fine structure also disappeared, which persisted longer than the absence of SOAEs. Interestingly, the thresholds at the maxima initially decreased which is an increased sensitivity (Long and Tubis, 1988b). During Aspirin intake, it is not surprising that the overall hearing level is then temporally raised (McFadden et al., 1984a, b).

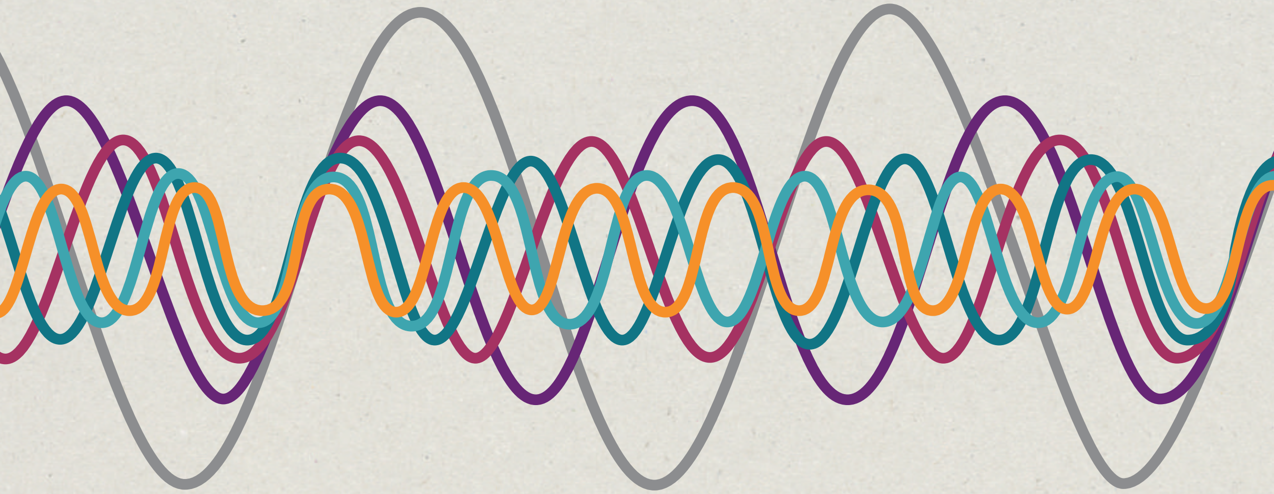
Another ototoxic drug, Carbamazepine, is applied to treat epilepsy and trigeminal neurologica, causing reversible shifts of the SOAE frequencies up to 4.5% (de Kleine et al., 2022). Interestingly the same drug produces a temporal downward shift in absolute pitch perception (Chaloupka et al., 1994). An increase of the stiffness of the cochlear partition may explain why SOAEs increase while perceived pitch decreases.

These two examples point out how ototoxic drugs may influence psychoacoustic outcomes. Drug consumption can thereby change the properties of SOAEs and tone perception. Such side effects would make intermediate measurements with an intake of different drug doses not comparable. Due to the potential effect of drugs on the results of psychoacoustic measurements, participants should have a clear and stable history of medication along the testing period.

5.5 Conclusions

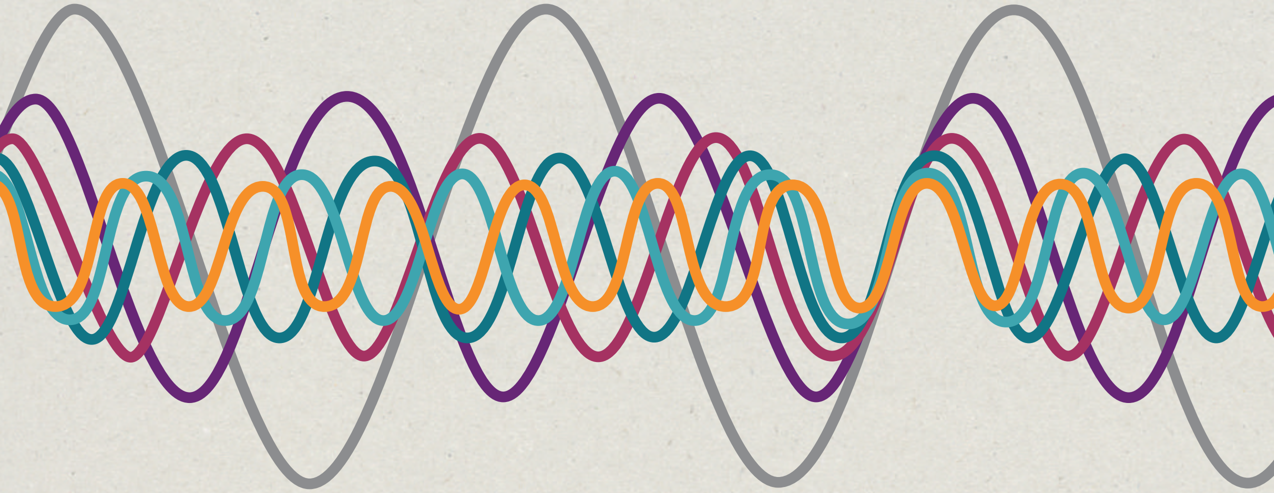
SOAEs are commonly present in normal hearing participants, often without any awareness by the researcher. SOAEs and their interactions can, however, become audible to the emitter. Complex interactions between the SOAE and the presented low-level pure tone stimuli may influence the results of psychoacoustic measurements, causing unexpected inter- and intra-participant variability. The experiments summarized in this review show that interactions between auditory stimuli and SOAE may explain a number of perceptual effects. This was found for non-complex stimuli (pure tones or noise) at relatively low stimulus levels.

The awareness of SOAE interactions is of importance especially when performing psychoacoustic measurements with tones close to perception threshold. The recognition of SOAEs and their interactions is important to choose suitable measurement parameters (e.g.: a probe tone frequency) and for the correct data interpretation. For low-level stimuli, interactions with SOAEs may add to variability in the outcome of a psychoacoustic experiment. Ideally, participants in psychoacoustic measurements (especially when using low level pure tones) should be screened for the presence of SOAEs, in order to optimize the test routine and identify possible interaction with SOAEs. Constant measurement conditions, such as body position and drug status, may be considered to obtain comparable results for each participant.



Chapter 6

General discussion



Preface

Even though the cochlea has already been studied for more than a hundred years (von Helmholz, 1874), many functional principles of its working are still not fully understood. One of these principles is the frequency selectivity (or frequency tuning) of the auditory system. Within the framework of this thesis, the frequency selectivity was investigated, aiming to contribute to a better understanding.

Various methods, which involve different levels of auditory processing, can be used to probe the frequency tuning of an individual's auditory system. In this way, the selectivity can be accessed at the peripheral level of the cochlea (suppression tuning curves, STCs), at the level of signal transmission with neural recordings (neural tuning curves, NTCs), and *all the way up* to the central stages of actual acoustic perception (psychoacoustical tuning curves, PTCs). Comparing the results of such methods enables the comparison of frequency selectivity of different processing levels.

However, not all of these methods are applicable in human research, as, for example, NTCs require single-unit neural recordings. Therefore, recordings in animals are used to complete the picture of frequency tuning at all these auditory processing levels. Moreover, by applying methods in animal models that are also feasible in human research, frequency tuning within the same level can be compared across species.

In this thesis, different methods were applied to test frequency selectivity at different levels of auditory processing, in humans and barn owls. Throughout different methods of testing frequency tuning, tuning curves (TCs) are typically V-shaped. As a measure of comparison between the different methods the sharpness of the TC-tip is used, where the threshold is lowest. The TC sharpness is commonly expressed by dividing the tip-frequency of the curve by its width 10 dB above the tip (Q_{10dB}). In the following paragraphs, three methods, from peripheral up to central auditory processing levels, will be explained and their application in measurements of frequency tuning will be discussed.

6.1 Summary

This thesis investigated the characteristics of spontaneous otoacoustic emissions (SOAEs) and their application in testing frequency selectivity. Studies were performed in barn owls and humans, probably two species with different mechanisms of SOAE generation. This inter-species comparison aimed to increase our understanding of (1) frequency selectivity differences between different vertebrate classes and (2) the comparability of different frequency tuning measurements of which not all are applicable in human research.

In the barn owl, it was shown that NTCs are sharper than STCs (Chapter 2). Contrary to findings in humans, larger SOAE amplitudes were not related to smaller peak widths. Narrower SOAE peaks had lower suppression thresholds, in barn owls. Internal noise in the SOAE generator could be responsible for wider SOAE amplitudes. Assuming that the suppression of an SOAE is the result of synchronization of the emission to the external tone, internal noise increases the synchronization/suppression threshold. According to this assumption, the internal noise acts contrary to synchronization/suppression to an external tone, that is why SOAEs containing less internal noise (narrow peaks) had lower suppression thresholds. Thus, the findings indicate different SOAE generation mechanisms between barn owls and humans. Whether a different SOAE generator may produce weak, noise-like SOAEs in the barn owl or/and the internal noise itself favors weak SOAEs that are less sensitive to suppression, remains speculative.

In barn owls, the comparison of NTCs and STCs revealed sharper neural tuning. A possible explanation why neural tuning appears to be sharper compared to cochlear tuning lies in the specificity of the method itself. Each afferent nerve fiber innervates one hair cell and thus a single spot on the BM. Consequently, the NTC reflects the frequency selectivity of that specific spot. In contrast, STCs and PTCs reflect interactions between the SOAE/probe tone and the suppressor/masker. These interactions take place along a certain portion of the basilar membrane (BM) and consequently involve tonotopic locations that are tuned to a range of frequencies. Another important fact to consider in such a comparison is that NTCs investigate the afferent system exclusively, transmitting information towards the brain, whereas PTCs and STCs may be influenced by the efferent system.

It was also investigated how language background may contribute to individual tuning differences in humans (Chapter 3). Previous research has shown that people with native tonal language acquisition often outperform others in psychoacoustic tasks (Pfordresher and Brown, 2009; Hove et al., 2010; Giuliano et al, 2011; Deutsch et al., 2004). To evaluate whether this outperformance could also be measured at the peripheral level of the cochlea, STCs of Asian tonal language (TL) native speakers were compared to those of Caucasian non-tonal language (NTL) native speakers. No significant difference in frequency tuning between TL and NTL speakers was found, suggesting that the behavioral outperformance of Asians with TL background is not based on enhanced cochlear tuning. Thus, the perceptual outperformance of TL speakers is presumably a result of the incidental acoustical training they receive by their language environment.

How objective SOAE measurements of frequency selectivity relate to behavioral approaches was evaluated in Chapter 4. In PTCs, the limit of frequency selectivity becomes obvious when the probe tone can no longer be separated from the masker. This observation is very similar to the measurement of frequency selectivity by using

STCs, where the lowest suppression threshold is measured when the external suppressor tone is of nearly the same frequency as the SOAE. This similarity between both measurements is a strong argument that the ability to selectively *perceive* sound depends on the frequency tuning that takes place at the level of the cochlea. It is believed that during SOAE suppression and tone masking the suppressor/masker excites the same auditory filter on the BM as the SOAE/probe tone.

Frequency selectivity in participants with SOAEs did not differ from participants without SOAEs. Therefore, the presence of SOAEs does not per se influence frequency selectivity. Interestingly, STCs in humans often show additional secondary suppression-dips (side-lobes) at the higher frequency flank (Manley and van Dijk, 2016; Chapters 3, 4). When evaluating PTCs at emission frequency of individuals with or without STC side-lobes, no significant differences in frequency selectivity between the groups were seen.

In order to measure PTCs, an individual needs to be trained to perform the tasks. The amount of training and experience as well as physical status (e.g., fatigue and concentration) can influence the test results. Moreover, in psychoacoustic tasks the information is centrally processed. Individuals may simply not be able to perceive the full extent of the sensed information, as attention influences our perception. For example, a person does not constantly perceive its own breathing noises. Thus at a perceptual level not all of the input may be available, which is the main principle of processing stimuli - the focus is on information that seems to be important. The world would be overwhelming if we would actually perceive everything that our receptors, auditory and others, detect.

The *McGurk effect* is an example of how speech perception can be influenced. This effect occurs when the mouth movements of the speaker conflict with the auditory cues. Participants see a person articulating a consonant whereby they record a voice articulating another consonant. Participants recognize the said acoustic speech signal when presented alone, but mis-categorize the speech signal when it is dubbed with incongruent visual speech. This effect is also described as an audiovisual illusion that presents a striking demonstration of multisensory integration between the auditory and the visual system.

Simply, the presence of SOAEs can lead to inter-participant variation in psychoacoustic tasks (Chapter 5). The interaction of SOAEs with presented low level pure tones may influence the performance of an individual drastically.

6.2 Implications

STCs can be used, in all individuals who have SOAEs, to evaluate the frequency selectivity of the cochlea, at the most peripheral level of auditory processing. They can be used to compare frequency tuning across individuals, and also across species, which is why SOAEs are of interest in diverse research questions.

The application of SOAE-STCs is especially useful when measuring cochlear tuning objectively. Since participants do not need to undergo any training to participate in STC measurements, it is also a time saving technique. Requirements of this method are the presence of measurable SOAEs and ideally a sound-attenuating chamber.

Comparing different methods of frequency tuning between species (Chapter 2) allows to formulate reasoned hypotheses. While the mammalian traveling wave is sharply tuned, the avian traveling wave appears to be rather broad (Xia et al., 2016). Such species-specific differences indicate strongly diverse mechanisms of inner ear tuning along vertebrates. The broad tuning on the BM may imply broader neural tuning, but single-unit recordings of the auditory nerve of the barn owl show the opposite (Köppl, 1997a; Fontaine et al., 2014). By comparing different methods, data gaps for human neural tuning can be filled by reasonable assumptions, based on studies that compare neural tuning of different vertebrate classes (e.g. Schoffelen et al., 2008).

In the cochlear model of Epp et al. (2015), the BM-oscillation that corresponds to a single SOAE frequency consists of a main peak of activity near the tonotopic location of the SOAE frequency, and an additional standing wave between the base of the cochlea and that tonotopic location. The standing waves have antinodes at BM-locations with characteristic frequencies that are approximately 0.5 and 1 octave above the emission frequency. Manley and van Dijk (2016) suggested that the STC side-lobes are due to interaction between the standing wave and the suppressor tone at the antinode location on the BM. Similarly, one might expect that a PTC would display side-lobes, if the probe tone is positioned at the frequency of an SOAE. The presence of suppression side-lobes would be consistent with the presence of standing waves on the BM.

The matter is, however, not so straightforward. In humans, STC side-lobes are not clearly related to PTC irregularities (Chapter 4; Engler et al., 2022). In addition, species that do not have traveling waves along the BM, as lizards, may still have side-lobes in STCs (e.g.: bobtail lizard: Manley et al., 1989; Köppl and Manley, 1994) and NTCs (Manley et a., 1988) measurements. Barn owls, however, lack high frequency side-lobes in STCs (Chapter 2; Engler et al., 2020b) and NTCs (Köppl, 1997a).

Suppression side-lobes are not present in every STC. In some cases, STC side-lobes are close to neighboring SOAEs. When these SOAEs interact, suppressing a neighboring SOAE may influence the SOAE of interest. Moreover, several common psychoacoustic measurements, such as pure tone audiometry, can be influenced by SOAE interactions (reviewed in Chapter 5). In the review chapter, some of such quite complex SOAE interactions are described, aiming to reach clarification, especially in unusual and variant test results. To allow a critical data interpretation, SOAE interactions that are likely to occur, need to be considered.

Presumably, acoustic experience can enhance the performance in psychoacoustical tasks, such as pitch discrimination. Acoustical experience can be shaped for example by musical training (e.g. Magne et al., 2006; Schön et al., 2004) or language experience (Pfordresher and Brown, 2009; Hove et al., 2010; Giuliano et al, 2011; Deutsch et al., 2004). The similar cochlear tuning of NTL and TL native speakers strongly indicates that language acquisition influences psychoacoustic performance. Frequency selectivity is also not per se enhanced in people with SOAEs compared to participants without SOAEs.

6.3 Future perspectives

Three methods were compared - assessing the frequency tuning at cochlear, neural, and central processing stages. Different models succeed in modeling specific characteristics of the SOAEs in different species. A comprehensive model that gives an explanation to all SOAE characteristics within one species is not available yet. The inter-species comparability needs further investigation. Developing models for different vertebrate classes may be the key to understand the fundamental principles and develop specific hypotheses of further processing.

Since musical training can influence psychoacoustic tasks (Magne et al., 2006; Schön et al., 2004), an effect that may be even larger in TL native speakers (Wong et al., 2007), the participants' level of musical training and language background should be taken into account when performing psychoacoustic measurements in general. Thus, for future measurements it is generally advised to state the individual's acoustical experience, whether that is language background or musical training, or both. Elaborating the data collection of people with different heritage and language acquisition will give some clarification to the influence of auditory training effects on psychoacoustic tasks. For example, PTCs of participants of Asian heritage without TL knowledge and musical training could be measured and compared to the data of Asian TL native speakers.

A further research study in Namibia was planned which aimed to investigate the cochlear tuning of native speakers using click-language, an idea by prof. G. A. Manley (University Oldenburg, Germany). We questioned whether (1) SOAEs, (2) cochlear tuning, and (3) audiograms are different compared to those of other language speakers. Unfortunately, the questions (for now) remain unanswered, as we were not able to perform the study during the pandemic crisis.

When testing frequency selectivity with a probe tone at an SOAE frequency, PTCs are similar to those of not spontaneous emitting participants. Presenting the probe tone at SOAE frequency does not give perceptual benefits (Chapter 4). The interaction of an SOAE with an external tone may be perceived by the emitter, influencing test results (reviewed in Chapter 5). So far, the perception of such interactions seems to influence acoustic measurements at low stimulus levels of non-complex stimuli only. When the masker is close to or at the frequency of the probe tone, masking is most effective and thus masking levels are low - a fact that is known for more than hundred years (Mayer, 1894; Wegel and Lane 1924).

To evaluate the degree of the interaction between an SOAE and external tones, PTCs of participants with SOAEs could be measured (1) just off SOAE frequency and (2) well off the emission frequency. While performing the task, participants should indicate whenever they perceive some frequency beating. The results of such a study may give some indications when interactions are perceived and to which degree they influence the PTC shape.

Especially when inter-participant variation is unusually high, SOAE screening is indicated to be aware of such potential interactions in the first place. Moreover, undesired SOAE interactions can be avoided by choosing suitable parameters.

6.4 Main findings

Frequency selectivity of barn owls

- NTCs are sharper tuned then STCs
- smaller SOAE amplitudes have higher suppression thresholds
- STCs have no side-lobes (contrary to human STCs)

Cochlear tuning across participants with different language background

- NTL and TL native speakers do not differ significantly in cochlear tuning
- SOAEs with center frequencies above 4.5 kHz are recorded in female Asian TL native speakers only, correlating with better high-frequency tone detection thresholds

STC-PTC comparison in humans

- STC side-lobes are not always seen in PTC shapes
- PTC irregularities occurred also in participants without SOAEs
- frequency selectivity is similar between participants with and without SOAEs
- STC tuning is significantly sharper compared to PTC selectivity, in participants without STC side-lobes

References

Abdala, C., Keefe, D.H., Oba, S.I., 2007. Distortion product otoacoustic emission suppression tuning and acoustic admittance in human infants: birth through 6 months. J. Acoust. Soc. Am. 121 (6), 3617–3627.

Ashmore, J.F., 1987. A fast motile response in guinea-pig outer hair cells: The cellular basis of the cochlear amplifier. J. Physiol. 388, 323-347.

Ashmore, J.F., 2008. Cochlear hair cell motility. Physiol. Rev. 88, 173-210.

Ashmore, J.F., Avan, P., Brownell, W.E., Dallos, P., Dierkes, K., Fettiplace, R., Grosh, K., Hackney, C.M., Hudspeth, A.J., Jülicher, F., Lindner, B., Martin, P., Meaud, J., Petit, C., Santos Sacchi, J.R., Canlon, B., 2010. The remarkable cochlear amplifier. Hear. Res. 266 (1-2), 1-17.

Avan, P., Büki, B., Maat, B., Dordain, M., Wit, H.P., 2000. Middle ear influence on otoacoustic emissions. I: noninvasive investigation of the human transmission apparatus and comparison with model results. Hear. Res. 120, 189–201.

Baiduc, R.R., Lee, J., Dhar, S., 2014. Spontaneous otoacoustic emissions, threshold microstructure, and psychophysical tuning over a wide frequency range in humans. J. Acoust. Soc. Am. 135 (1).

Bartsch, T.F., Hengel, F.E., Oswald, A., Dionne, G., Chipendo, I.V., Mangat, S.S., Shatanfoy, M.E., Shapiro, L., Müller, U., Hudspeth, A.J., 2019. Elasticity of individual protocadherin 15 molecules implicates tip links as the gating springs for hearing. Proc. Natl. Acad. Sci. 116 (22), 11048-11056.

Bell, A., 1992. Circadian and menstrual rhythms in frequency variations of spontaneous otoacoustic emissions from human ears. Hear. Res. 58, 91-100.

Bent, T., Bradlow, A.R., Wright, B.A., 2006. The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. J. Exp. Psychol. Hum. 32 (1), 97–103.

Bergevin, C. and Shera, C.A., 2010. Coherent reflection without traveling waves: on the origin of long-latency otoacoustic emissions in lizards. J. Acoust. Soc. Am. 127 (4), 2398-2409.

Bergevin, C., Freeman, D.M., Saunders, J.C., Shera, C.A., 2008. Otoacoustic emissions in humans, birds, lizards, and frogs: evidence for multiple generation mechanisms. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 194 (7),665-683.

Bergevin, C., Manley, A.G., Köppl, C., 2015. Salient features of otoacoustic emissions are common across tetrapod groups and suggest shared properties of generation mechanisms. Proc. Natl. Acad. Sci. 112 (11), 3362-3367.

Berglund, A.M. and Ryugo, D.K., 1987. Hair cell innervation by spiral ganglion neurons in the mouse. J. Comp. Neuro. 255, 560-570.

Beurg, M.B., Tan, X., Fettiplace, R., 2013. A prestin motor in chicken auditory hair cells: active force generation in a nonmammalian species. Neuron 79, 69-81.

Bidelman, G.M., Hutka, S., Moreno, S., 2013. Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. PLoS One 8 (4), 1–11.

Bialek, W. and Wit, H.P., 1984. Quantum limits to oscillator stability: Theory and experiments on acoustic emissions from the human ear. Phys. Lett. 104 A (3), 173-178.

Bilger, R.C., Matthies, M.L., Hammel, D.R., 1990. Genetic implications of gender differences in the prevalence of spontaneous otoacoustic emissions. J. Speech Hear. Res 33, 418-432.

Braun, B. and Johnson, E.K., 2011. Question or tone 2? How language experience and linguistic function guide pitch processing. J. Phon. 39, 585–594.

Bright, K.E., 1985. Microstructure audiograms and psychophysical tuning curves from subjects with spontaneous otoacoustic emissions, Dissertation, The University of Arizona, Tucson, AZ.

Brown, M.C., 1987. Morphology of labeled afferent fibers in the guinea pig cochlea. J. Comp. Neuro. 260, 591-604.

Brownell, W.E., 1990. Outer hair cell electromotility and otoacoustic emissions. Ear Hear. 11 (2), 82-92.

Brownell, W.E., Bader, C.R., Bertrand, D., de Ribaupierre, Y., 1985. Evoked mechanical response of isolated outer hair cells. Science 227 (4683), 194-196.

Büki B., de Kleine, E., Wit, H.P., Avan, P. 2002. Detection of intracochlear and intracranial pressure changes with otoacoustic emissions: a gerbil model. Hear Res. 167 (1-2), 180-91.

Burns, E.M., 2009. Long-term stability of spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 125 (5), 3166 -3176.

Burns, E.M., 2017. Even-longer-term stability of spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 142 (4), 1828 - 1831.

Burns, E.M., Strickland, E.A., Tubis, A., Jones, K., 1984. Interactions among spontaneous otoacoustic emissions. I. Distortion products and linked emissions. Hear. Res. 16, 217–278.

Chaloupka, V., Mitchell, S., Muirhead, R., 1994. Observation of a reversible, medication-induced change in pitch perception. J. Acoust. Soc. Am. 96 (1), 145-149.

Chan, J.C.Y. and McPherson, B., 2001. Spontaneous and transient evoked otoacoustic emissions: a racial comparison. J. Audiol. Med. 10 (1), 20–32.

Charaziak, K.K. and Siegel, J.H., 2014. Estimating cochlear frequency selectivity with stimulus frequency otoacoustic emissions in chinchillas. J. Assoc. Res. Otolaryngol. 15, 883–896.

Charaziak, K.K., Souza, P., Siegel, J.H., 2013. Stimulus-frequency otoacoustic emission suppression tuning in humans: Comparison to behavioral tuning," J. Assoc. Res. Otolaryngol. 14, 843–862.

Chen, L., Sun, W., Salvi, R.J., 2001. Electrically evoked otoacoustic emissions from the chicken ear. Hear. Res. 161 (1-2), 54-64.

Ciuman, R.R., 2010. The Efferent System or Olivocochlear Function Bundle – Fine Regulator and Protector of Hearing Perception. Int. J. Biomed. Sci. 6 (4), 276-288.

Corwin, J.T. and Cotanche, D.A., 1988. Regeneration of hair cells after acoustic trauma. Science 240, 1772-1774.

Dallos, P., 1992. The active cochlea. J. Neurosc. 12 (12), 4575-4585.

Dallos, P., 1996. Overview: cochlear neurobiology. In: Dallos, P., Popper, A.N., Fay, R.R. (Eds.), The Cochlea. Springer Verlag, New York, 1-43.

Dallos, P., Xudong, W., Cheatham, M.A., Gao, J., Zheng, J., Anderson, C.T., Jia, S., Wang, X., Cheng, W.H.Y., Sengupta, S., He, D.Z.Z., Zuo, J., 2008. Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. Neuron 58 (3), 333-339.

Davis, H., 1983. An active process in the cochlear mechanics. Hear. Res. 9, 79-90.

de Kleine, E., Wit, H.P., van Dijk, P., Avan, P., 2000. The behavior of spontaneous otoacoustic emissions during and after postural changes. J. Acoust. Soc. Am. 107, 3308-3316.

de Kleine, E., Maat, B., Metzemaekers, J.D.M., van Dijk, P., 2022. Carbamazepine induces upward frequency shifts of spontaneous otoacoustic emissions. Associ. Res. submitted

Deutsch, D., Henthorn, T., Dolson, M., 2004. Absolute pitch, speech, and tone language: some experiments and a proposed framework. Music Percept. 21 (3), 339–356.

Deutsch, D., Henthorn, T., Marvin, E., Xu, H.S., 2006. Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. J. Acoust. Soc. Am. 119 (2), 719–722.

Dooling, R.J., 1982. Auditory perception in birds. In: Kroodsma, D.E., Miller, E.H. (Eds.), Acoustic Communication in Birds (1). Academic Press, New York, pp. 95-130.

Dyson, M.L., Klump, G.M., Gauger, G., 1998. Absolute hearing thresholds and critical masking ratios in the European barn owl: a comparison with other owls. J. Comp. Physiol. A. 182, 695-702.

Elliott, E., 1958. A ripple effect in the audiogram. Nature 181, 1076.

Engler, S., de Kleine, E., Avan, P., van Dijk, P. 2020a. Frequency selectivity of tonal language native speakers probed by suppression tuning curves of spontaneous otoacoustic emissions. Hear. Res. 298 (108100), 1-8.

Engler, S., Köppl, C., Manley, G.A., de Kleine, E., van Dijk, P., 2020b. Suppression tuning of spontaneous otoacoustic emissions in the barn owl (Tyto alba). Hear. Res. 385, 1–9.

Engler, S., Gaudrain, E., de Kleine, E., van Dijk, P., 2022. Relationship between irregularities in spontaneous otoacoustic emissions suppression and psychophysical tuning curves. J. Acoust. Soc. Am. 151 (2), 1055-1063.

Epp, B., Verhey, J.L., Mauermann, M., 2010. Modeling cochlear dynamics: Interrelation between cochlea mechanics and psychoacoustics. J. Acoust. Soc. Am. 128, 1870-1883.

Epp, B., Wit, H., van Dijk, P., 2015. Clustering of cochlear oscillations in frequency plateaus as a tool to investigate SOAE generation. In: Mechanics of Hearing: Protein to Perception, Proceedings of the 12th International Workshop on the Mechanics of Hearing, AIP Conf. Proc., 1703, Karavitaki, K.D., Corey, D.P. (Eds.), p. 090025.

Epp, B., Manley, G.A., van Dijk, P., 2018. The mechanisms underlying multiple lobes in SOAE suppression tuning curves in a transmission line model of the cochlea. Am. Institute of Physics 1-6, 090005.

Fischer, F.P., 1992. Quantitative analysis of the innervation of the chicken basilar papilla. Hear. Res. 61, 167-178.

Fischer, F.P., Köppl, C., Manley, G.A., 1988. The basilar papilla of the barn owl Tyto alba: a quantitative morphological SEM analysis. Hear. Res. 34, 87-102.

Flottorp, G., 1953. Pure-tone tinnitus evoked by acoustic stimulation: the idiophonic effect. Acta. Oto-Laryngologica 43 (4-5), 396-415.

Fontaine, B., Köppl, C., Pena, J.L., 2014. Reverse correlation analysis of auditory-nerve fiber responses to broadband noise in a bird, the barn owl. J. Assoc. Res. Otolaryngol. 16, 101-119.

Frick, L.R. and Matthies, M.L., 1988. Effects of external stimuli on spontaneous otoacoustic emissions. Ear and Hear. 9 (4), 190-197.

Gandour, J., 1998. Aphasia in tone languages. In: Coppens, P., Basso, A., Lebrun, Y. (Eds.), Aphasia in Atypical Populations. Lawrence Erlbaum, Hillsdale, NJ, pp. 117–141.

Gandour, J., Wong, D., Hutchins, G.D., 1998. Pitch processing in the human brain is influenced by language experience. Neuroreport 9 (2), 2115–2119.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., van Lancker, D., Hutchins, G.D., 2000. A crosslinguistic PET study of tone perception. J. Cogn. Neurosci. 12 (1), 207–222.

Gardner, M.B., Gardner, R.S., 1973. Problem of localization in the median plane: effect of pinnae cavity occlusion. J. Acoust. Soci. Am. 53, 400-408.

Geisler, C.D., Yates, G.K., Patuzzi, R.B., Johnstone, B.M., 1990. Saturation of outer hair cell receptor currents causes two-tone suppression. Hear. Res. 44, 241-256.

Giuliano, R.J., Pfordresher, P.Q., Stanley, E.M., Narayana, S., Wicha, N.Y.Y., 2011. Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. Front. Psychol. 146 (2), 1–12.

Gold, T., 1948. Hearing II. The physical basis of the action of the cochlea. Proc. R. Soc. E. B 135, 492-498.

Gummer, W.G., Smolders, J.W.T., Klinke, R., 1987. Basilar membrane motion in the pigeon measured with the Mossbauer technique. Hear. Res. 29, 63-92.

He, D.Z.Z., Beisel, K.W., Chen, L., Ding, D.-L., Jia, S., Fritzsch, B., Salvi, R., 2003. Chick hair cells do not exhibit coltage-dependent somatic motility. J. Phyisiol. 546 (2), 511-520.

Heise, S.J., Verhey, J.L., Mauermann, M., 2008. Automatic screening and detection of threshold fine structure. Int. J. Audiol. 47, 520–532.

Heise, S. J., Mauermann, M., Verhey, J.L., 2009a. Investigating possible mechanisms behind the effect of threshold fine structure on amplitude modulation perception. J. Acoust. Soc. Am. 126 (5), 2490-2500.

Heise, S.J., Mauermann, M., Verhey, J.L., 2009b. Threshold fine structure affects amplitude modulation perception," J. Acoust. Soc. Am. 125 (1), 33–38.

Hirokawa, N., 1978. The ultrastructure of the basilar papilla of the chick. J. Comp. Neur. 181, 361-374.

Horst, J.W., Wit, H.P., Ritsma, R.J., 1983. Psychophysical aspects of cochlear acoustic emissions ("Kemp-tones"). Hearing, physiological bases and psychophysics, proceedings of the sixed international symposium, Klinke, R. and Hartmann, R. (Eds.), 1 (1), 89-96.

Hove, M.J., Sutherland, M.E., Krumhansl, C.L., 2010. Ethnicity effects in relative pitch learning. Psychon. Soc., Inc. 17 (3), 310–316.

Howard, J., Hudspeth, A.J., 1988. Compliance of hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell. Neuron 1, 189-199.

Hundspeth, A.J., 2008. Making an effort to listen: Mechanical amplification in the ear. Neuron 59, 530-545.

Hundspeth, A.J., 2014. Integrating the active process of hair cells with cochlear function. Nat. Rev. Neurosci. 15, 600-614.

Jones, K., Tubis, A., Long, G.R., Burns, E.M., Strickland, E.A., 1986. Interactions among multiple spontaneous otoacoustic emissions. In Peripheral Auditory Mechanisms, edited by Allen, J.B., Hall, J.L., Hubbard, A., Neely, S.T., Tubis, A. (Eds.), Springer Verlag, New York, pp. 266–273.

Joris, P.X., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., Shera, C.A., 2011. Frequency selectivity in old-world monkeys corroborates sharp cochlear tuning in humans. Proc. Nat. Acad. Sci. 108 (42), 17516–17520.

Kachar, B., Brownell, W.E., Altschuler, R., Fex, J., 1986. Electrokinetic shape changes of the cochlear outer hair cells. Nature 322, 365-368.

Kapadia, S., Lutman, M.E., 1999. Reduced 'audiogram ripple' in normally-hearing subjects with weak otoacoustic emissions. Audiol. 38, 257-261.

Kemp, D.T., 1978. Stimulated acoustic emissions from within the human auditory system. J. Acoust. Soc. Am. 64 (5), 1386–1391.

Kemp, D.T., 1979a. Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. Arch. Otorinolaryngol. 224, 37-47

Kemp, D.T., 1979b. The evoked cochlear mechanical response and auditory microstructure - evidence for a new element in cochlear mechanics. Scand. Audiol. Suppl. 9, 35–47.

Kemp, D.T., 1980. Towards a model for the origin of cochlear echoes. Hear. Res. 2, 533–548.

Kemp, D.T. and Chum, R.A., 1980. Observations on the generator mechanism of stimulus frequency acoustic emissions-two tone suppression," in Psychophysical, Physiological and Behavioral Studies in Hearing, Boer, E., Viergever, M.A. (Eds). Delft University Press, Delft, pp. 34–41.

Kemp, D.T., 2002. Otoacoustic emissions, their origin in cochlear function, and use. Br. Med. Bull. 63 (1), 223-2541.

Kiang, N.Y., Liberman, M.C., Levine, R.A., 1976. Auditory-nerve activity in cats exposed to ototoxic drugs and high-intensity sounds. Ann. Otol. Rhinol. Laryngol. 85, 752-768.

Konishi, M., 1973. How the owl tracks its prey. Am. Sci. 61, 414-424.

Köppl, C., 1995. Otoacoustic emissions as an indicator for active cochlear mechanics: a primitive property of vertebrate auditory organs. In: Manley, G.A., Klump, G.M., Koppl, C., Fastl, H., Oeckinghaus, H. (Eds.), Advances in Hearing Research. World Scientific, Singapore, pp. 207-218.

Köppl, C., 1997a. Frequency tuning and spontaneous activity in the auditory nerve and cochlear nucleus magnocellularis of the barn owl Tyto alba. J. Neurophysiol. 77, 364-377.

Köppl C., 1997b. Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba. J. Neurosci. 17 (9), 3312-3321.

Köppl C., 2011. Birds - same thing, but different? Convergent evolution in the avian and mammalian auditory systems provides informative comparative models. Hear. Res. 273, 65-71.

Köppl, C., 2015. Avian hearing. In: Scanes, C.G. (Ed.), Sturkie's Avian Physiology, sixth ed. Elsevier, pp. 71-111 (chapter 6).

Köppl, C. and Manley, G.A., 1993. Spontaneous otoacoustic emissions in the bobtail lizard. I: general characteristics. Hear. Res. 71, 157-169.

Köppl, C. and Manley, G.A., 1994. Spontaneous otoacoustic emissions in the bobtail lizard. II: interactions with external tones. Hear. Res. 72, 159-170.

Köppl, C., Gleich, O., Manley, G.A., 1993. An auditory fovea in the barn owl cochlea. J. Comp. Physiol. A. 171, 695-704.

Kraus, N. and Banai, K., 2007. Auditory-processing malleability. Focus on language and music. Assoc. Physiol. Sci. 16 (2), 105–110.

Krishnan, A., Xu, Y., Gandour, J., Cariani, P., 2005. Encoding of pitch in the human brainstem is sensitive to language experience. Cogn. Brain Res. 25, 161–168.7

Krishnan, A., Gandour, J.T., 2009. The role of the auditory brainstem in processing linguistically-relevant pitch patterns. Brain Lang. 110 (3), 135–148.

Krishnan, A., Gandour, J.T., Bidelman, G.M., 2010a. Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. Brain Res. 1313, 124–133.

Krishnan, A., Gandour, J.T., Bidelman, G.M., 2010b. The effects of tone language ex- perience on pitch processing in the brainstem. J. Neurolinguist. 23 (1), 81–95.

Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. Nat. Rev. Neurosci. 5, 831–843.

Krumm, B., Klump, G., Koppl, C., Langemann, U., 2017. Barn owls have ageless ears. Proc. R. Soc. B 284, 1-7.

Langemann, U., Hamann, I., Friebe, A., 1999. A behavioral test of presbycusis in the bird auditory system. Hear. Res. 137, 68-76.

Langers, D.R.M. and van Dik, P., 2012. Auditory cortex with minimally salient acoustic stimulation. Cerebral Correx 22, 2024–2038.

Larsen, O.N., Christensen-Dalsgaars, J., Jensen, K.K., 2016. Role of intracranial cavities in avian directional hearing. Biol. Cybern. 110 (4-5), 319-331.

Liberman, M.C. and Guinan, J.J., 1998. Feedback control of the auditory periphery: Anti-masking effects of middle ear muscles vs. olivocochlear efferents. J. Commun. Disord, 31, 471-483.

Liberman, M.C., Dodds, L.W., Pierce, S., 1990. Afferent and efferent innervation of the cat cochlea. Quantitative analysis with light electron microscopy. J. Comp. Neurol. 301, 443-460.

Liberman, M.C., Gao, J., He, D.Z.Z., 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.

Long, G., 1998. Perceptual consequences of the interactions between spontaneous otoacoustic emissions and external tones. I. Monaural diplacusis and aftertones. Hear. Res. 119, 49-60.

Long, G.R., 1993. Perceptual consequences of otoacoustic emissions. In: Schick, A. (Eds.), Results of the sixth Oldenburg symposium on psychological acoustics, 59-79.

Long, G.R., and Tubis, A., 1988a. Investigations into the nature of the association between threshold microstructure and otoacoustic emissions," Hear. Res. 36, 125–138.

Long, G. R., and Tubis, A., 1988b. Modification of spontaneous and evoked otoacoustic emissions and associated psychoacoustic microstruc- ture by aspirin consumption," J. Acoust. Soc. Am. 84 (4), 1343–1353.

Long, G.R., Tubis, A., Jones, K.L., 1991. Modeling synchronization and suppression of spontaneous otoacoustic emissions using Van der Pol oscillators: Effects of aspirin administration. J. Acoust. Soc. Am. 89, 1201-1212.

Macrane, J.H., 1972. Effects of body position on the auditory system. J. Speech Hear. Res. 15 (2), 330-339.

Magne, C., Schön, D., Besson, M., 2006. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. J. Cogn. Neurosci. 18 (2), 199–211.

Manley, G.A., 2000. Cochlear mechanisms from a phylogenetic viewpoint. Proc. Natl. Acad. Sci. U.S.A. 97, 11736-11743.

Manley, G.A., 2001. Evidence for an active process and a cochlear amplifier in non- mammals. J. Neurophysiol. 86, 541-549.

Manley, G.A., 2004. Spontaneous otoacoustic emissions in monitor lizards. Hear. Res. 189, 41-57.

Manley, G.A., 2006. Spontaneous otoacoustic emissions from free-standing stereovillar bundles of ten species of lizard with small papillae. Hear. Res. 212, 33-47.

Manley, G.A., 2017. The mammalian Cretaceous cochlear revolution. Hear. Res. 352, 23-29.

Manley, G.A. and Gleich, O., 1992. Evolution and specialization of function in the avian auditory periphery. The Evolutionary Biology of Hearing, pp. 561-580.

Manley, G.A. and Köppl C., 1998. Phylogenetic development of the cochlea and its innervation. Curr. Opin. Neurobiol. 8, 468-474.

Manley, G.A. and Köppl, C., 2008. What have lizard ears taught us about auditory physiology? Hear. Res. 238, 3-11.

Manley, G.A. and Taschenberger, G., 1993. Spontaneous otoacoustic emissions from a bird: a preliminary report. In: Duifhuis, H., Horst, J.W., van Dijk, P., van Netten, S.M. (Eds.), Biophysics of Hair Cell Sensory Systems. World Scientific Publishing Co., Singapore, pp. 33-39.

Manley, G.A. and van Dijk, P., 2008. Otoacoustic emissions in amphibians, lepidosaurs and archosaurs. In: Manley, G.A., Fay, R.R., Popper, A. (Eds.), Active Processes and Otoacoustic Emissions in Hearing; Springer Handbook of Auditory Research, vol. 30. Springer-Verlag, New York, ISBN 978-0-387-71467-7, pp. 211-260.

Manley, G.A. and van Dijk, P., 2016. Frequency selectivity of the human cochlea. Hear. Res. 336, 53-62.

Manley, G.A., Graeme, K.Y., Köppl, C., 1988. Auditory peripheral tuning: evidence for a simple resonance phenomenon in the lizard Tiliqua. Hear. Res. 33, 181-190.

Manley, G.A., Köppl, C., Johnstone, B.M., 1990. Peripheral auditory processing in the bobtail lizard Tiliqua rugosa. J. Comp. Physiol. 167, 89-99.

Manley, G.A., Gallo, L., Köppl, C., 1996. Spontaneous otoacoustic emissions in two gecko species, Gekko gecko and Eublepharis macularius. J. Acoust. Soc. Am. 99, 1588-1603.

Manley, G.A., Taschenberger, G., Oeckinghaus, H., 1999. Influence of contralateral acoustic stimulation on distortion-product and spontaneous otoacoustic emissions in the barn owl. Hear. Res. 138, 1-12.

Manley, G.A., Köppl, C., Bergevin, C., 2015. Common Substructure in otoacoustic emission spectra of land vertebrates. In: Karavitaki, K.D., Corey, D.P. (Eds.), Mechanics of Hearing: Protein to Perception, American Institute of Physics, vol. 1703. AIP Conference Proceedings, Melville, NY, 090012/1-5.

Martin, G.K., Lonsbury-Martin, B.L., Probst, R., Coats, A.C., 1988. Spontaneous otoacoustic emissions in a nonhuman primate. I. Basic features and relations to other emissions. Hear. Res. 33, 49-68.

Mauermann, M., Long, G.R., Kollmeier, B.,2004. Fine structure of hearing threshold and loudness perception. J. Acoust. Soc. Am. 116 (2), 1066–1080.

McFadden, D. and Mishra, R., 1993. On the relation between hearing sensitivity and otoacoustic emissions. Hear. Res. 71, 208-213.

McFadden, D. and Plattsmier, H.S., Pasanen, E.G., 1984a. Aspirin-induced hearing loss as a model of sensorineural hearing loss. Hear. Res. 16, 251-260.

McFadden, D., Plattsmier, H.S., Pasanen, E.G., 1984b. Temporary hearing loss induced by combinations of intense sounds and nonsteroidal anti-inflammatory drugs. A. J. Ototlaryngol. 5, 236-241.

McFadden, D., Pasanen, E.G., Maloney, M.M., Leshikar, E.M., Pho, M.H., 2018. Correlations between otoacoustic emissions and performance on common psychoacoustical tasks. J. Acoust. Soc. Am. 143 (4), 2355-2367.

McKetton, L., Purcell, D., Stone, V., Grahn, J., Bergevin, C., 2018. No otoacoustic evidence for a peripheral basis of absolute pitch. Her. Res. 370, 201–208.

Micheyl, C., and Collet, L., 1994. Interrelations between psychoacoustical tuning curves and spontaneous and evoked otoacoustic emissions. Scand. Audiol. 23, 171–178.

Micheyl, C., Delhommeau, K., Perrot, X., Oxenham, A.J., 2006. Influence of musical and physioacoustical training on pitch discrimination. Hear. Res. 219, 36–47.

Moerel, M., de Martino, F., Formisano, E., 2012. Processing of neural sounds in human auditory cortex: Tonotopy, spectral tuning, and relation to voice sensitivity. J. Neurosci. 32, 14205–14216.

Moiseff, A. and Konishi, M., 1981. The owl's interaural pathway is not involved in sound localization. J. Comp. Physiol. 144, 299-304.

Moore, B.C.J., 1978. Psychophysical tuning curves measured in simultaneous and forward masking. J. Acoust. Soci. Am. 63, 524–532.

Moore, B.C.J., 1989. An Introduction to the Psychology of Hearing, Moore, B.C.J. (Eds.). Academic Press, Cambridge, MA.

Moore, B.C.J., 1995. Frequency analysis and masking. In: Handbook of Perception and Cognition, 2nd ed., Moore, B.C.J. (Eds.). Academic Press, Cambridge, MA, pp. 161–205.

Moulin, A., Collet, L., Veuillet, E., Morgon, A., 1993. Interrelations between transiently evoked otoacoustic emissions, spontaneous otoacoustic emissions and acoustic distortion products in normally hearing subjects. Hear. Res. 65, 216-233.

Murphy, W.J., Tubis, A., Talmadge, C.L., and Long, G.R., 1995. Relaxation dynamics of spontaneous otoacoustic emissions perturbed by external tones. II. Suppression of interacting emissions. J. Acoust. Soc. Am. 97, 3711–3720.

Narayan, S.S., Temchin, A.N., Recio, A., Ruggero, M.A., 1998. Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. Science 282, 1882-1884.

Norrix, L.W., and Glattke, T.J., 1996. Distortion product otoacoustic emissions created through the interaction of spontaneous otoacoustic emissions and externally generated tones. J. Acoust. Soc. Am. 100, 945–955.

Ohyama, K., Wada, H., Kobayashi, T., Takasaka, T., 1991. Spontaneous otoacoustic emissions in the Guinea pig. Hear. Res. 56, 111-121.

Otte, J., Schuknecht, H.F., Kerr, A.G., 1978. Ganglion cell populations in normal and pathological human cochleae: implication for cochlear implantation. Laryngoscope 88, 1231-1246.

Oxenham, A.J. and Shera, C.A., 2003. Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J. Assoc. Res. Otolaryngol. 4, 541–554.

Palanca-Castan, N., Laumen, G., Reed, D., Köppl, C., 2016. The binaural interaction component in barn owl (Tyto alba) presents few differences to mammalian data. JARO 17, 577-589.

Palmer, A.R. and Wilson, J.P., 1982. Spontaneous and evoked acoustic emissions in the frog Rana esculenta. J. Physiol. 324, 66.

Pasanen, E.G. and McFadden, D., 2000. An automated procedure for identifying spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 108 (3), 1105-1116.

Patuzzi, R., 1996. Cochlear micromechanics and macromechanics. In: The cochlea Dallos, P., Popper, A.N., Fay, R.R. (Eds.). New York: Springer, 186-257.

Payne, R.S., 1971. Acoustic location of prey by barn owls (Tyto alba). J. Exp. Biol. 54, 535-573.

Peng, A.W. and Ricci, A.J., 2011. Somatic motility and hair bundle mechanics, are both necessary for cochlear amplification? Hear. Res. 273 (1-2), 109-122.

Penner, M.J. and Zhang, T., 1997. Prevalence of spontaneous otoacoustic emissions in adults revisited. Hear. Res. 103, 28-34.

Pfordresher, P.Q. and Brown, S., 2009. Enhanced production and perception of musical pitch in tone language speakers. Attent. Percept. Psychophys. 71 (6), 1385–1398.

Pickles, J.O., Comis, S.D., Osborne, M.P., 1984. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. Hear. Res. 15, 103-112.

Plack, C.J., 2018. The sense of hearing. Routledge, third edition.

Plassmann, W. and Brändle, K., 1992. A functional model of the auditory system in mammals and its evolutionary implications. In: Webster, D.B., Fay, R.R., Popper, A.N. (Eds.), The Evolutionary Biology of Hearing. Springer-Verlag, New York, pp. 637–653.

Probst, R., Lonsbury-Martin, B.L., Martin, G.K., 1991. A review of otoacoustic emissions. J. Acoust. Soc. Am. 89, 2027–2067.

Probst, R. and Harris, F.P., 1996. Effect of otoacoustic emissions on just-noticable differences for intensity in normal hearing subjects. Acoust. Soc. Am. 100 (1), 504-510.

Quine, D.B.and Konishi, M., 1974. Absolute frequency discrimination in the barn owl. J. comp. Physiol 93, 347-360.

Rabinowitz, W.M. and Widin, G.P., 1984. Interaction of spontaneous oto-acoustic emissions and external sounds. J. Acoust. Soc. Am. 76, 1713-1720.

Raphael, Y. and Altschuler, R.A., 2003. Brain Res. Bull., 397-422.

Rask-Andersen, H., Li, H., Loewenheim,H., Mueller, M., Pfaller, K., Schrott-Fischer, A., Glueckert, R., 2017. Supernumerary human hair cells—signs of regeneration or impaired development? A field emission scanning electron microscopy study. Upsala J. Med. Sci. 122 (1), 11-19.

Rhode, W.S., 1995. Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. J. Acoust. Soc. Am. 97 (4), 2414-2429.

Robles, L. and Ruggero, M.A., 2001. Mechanics of the mammalian cochlea. Physiol. Rev. 81 (3), 1305–1352.

Ruggero, M.A. and Rich, N.C., 1991. Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane. J. Neurosci. 11 (4), 1057-1067.

Ryals, B.M. and Rubel, E.W., 1988. Hair cell regeneration after acoustic trauma in adult Coturnix quail. Am. Assoc. Adv. Sci. 240 (4860), 1774-1776.

Ryals, B.M., Dent, M.L., Dooling, R.J., 2013. Return of function after hair cell regeneration. Hear. Res. 297, 113-120.

Salmelin, R., Schnitzler, A., Parkkonen, L., Biermann, K., Helenius, P., Kiviniemi, K., Kuukka, K., Schmitz, F., Freund, H., 1999. Native language, gender, and functional organization of the auditory cortex. Proc. Natl. Acad. Sci. 96, 10466-10465.

Schellenberg, E.G. and Trehub, S.E., 2008. Is there an Asian advantage for pitch memory? Music Percept. 25 (3), 241–252.

Schloth, E. and Zwicker, E., 1983. Mechanical and acoustical influences on spontaneous oto-acoustic emissions. Hear. Res. 11 (3), 285-293.

Schön, D., Magne, C., Besson, M., 2004. The music of speech: music training facilitates pitch processing in both music and language. Psychophysiology 41, 341–349.

Sellick, P.M., Patuzzi, R., Johnstone, B.M., 1982. Measurement of basilar membrane motion in the Guinea pig using the Mossbauer technique. J. Acoust. Soc. Am. 72 (1), 131-141.

Shahnaz, N., 2008. Transient evoked otoacoustic emissions (TEOAEs) in Caucasian and Chinese young adults. Int. J. Audiol. 47 (2), 76-83.

Shahnaz, N. and Davies, D., 2006. Standard and multifrequently tympanometric norms for Caucasian and Chinese young adults. Ear Hear. 75–90.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270 (5234), 303–304.

Shera, C.A., 2003. Mammalian spontaneous otoacoustic emissions are amplitude- stabilized cochlear standing waves. J. Acoust. Soc. Am. 113 (5), 2762-2772.

Shera, C.A., 2022. Whistling while it works: Spontaneous otoacoustic emissions and the cochlear amplifier. J. Acoust. Res. Otollaryngol., 17-25.

Shera, C.A., Guinan, J.J., Oxenham, A.J., 2002. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc. Natl. Acad. Sci. 99 (5), 3318–3323.

Shera, C.A., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., Joris, P.X., 2011. Otoacoustic estimates of cochlear tuning: testing predictions in macaque. In: Shera, C.A., Olson, E.S. (Eds.), What Fire Is in My Ears: Progress in Auditory Biomechanics, vol. 1403. American Institute of Physics, AIP Conf Proc., pp. 286-292

Smith, C.A., Konishi, M., Schuff, N., 1985. Structure of the arn owl's (*Tyto alba*) inner ear. Hear. Res. 17, 237-247.

Smolders, J.W.Th, 1999. Functional recovery in the avian ear after hair cell regeneration. Audiol. Neuro. Otol. 4, 286-302.

Smolders, J.M., Pfenningdorff, D., Klinke, R., 1995. A functional map of the pigeon basilar papilla: correlation of the properties of single auditory nerve fibres and their peripheral origin. Hear. Res. 92 (1-2), 151-169.

Smurzynski, J. and Probst, R., 1999. Intensity discrimination, temporal integration and gap detection by normally-hearing subjects with weak and strong otoacoustic emissions. Audiol. 38 (5), 251-256.

Smurzynski, J. and Probst, R., 2000. Detection of partially filled and empty temporal gaps in low-pass noise in subjects with and without spontaneous otoacoustic emissions (SOAEs). In: Wada, H., Takasaka, T., Ikeda, K., Ohyama, K., Koike, T., (eds.), Recent Developments in Auditory Mechanics. Singapore: World Scientific, 463–469.

Smurzynski, J., Harris, F.P., Probst, R., 2001. Otoacoustic emissions in relation to auditory perception: Beyond the pure-tone audiogram. Semin. Hear. 22 (4), 361-375.

Spoendlin, H.S., 1985. Anatomy of the cochlear innervation. Am. J. Otolaryngol. 6, 453-467.

Stratonovich, R.L., 1967. Topics in the Theory of Noise, vol. 3. Science publisher, pp. 222-227 (chapter 9).

Strickland, E.A., Burns, E.M., Tubis, A., 1985. Incidence of spontaneous otoacoustic emissions in children and infants. J. Acoust. Soc. Am. 78, 931-935.

Suga, N., Ma, X., Gao, E., Sakai, M., Chowdhury, S.A., 2003. Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists. Speech Commun. 41, 189–200.

Sul, B., Iwasa, K.H., 2009. Effectiveness of hair bundle mortality as the cochlear amplifier. Biophys. J. 97, 2653-2663.

Sumner, C.J., Wells, T.T., Bergevin, C., Sollini, J., Kreft, H.A., Palmer, A.R., Oxenham, A.J., Shera, C.A., 2018. Mammalian behaviour and physiology converge to confirm sharper cochlear tuning in humans. Proc. Natl. Acad. Sci. 115 (44), 11322–11326.

Swaminathan, J., Krishnan, A., Gandour, J.T., 2008. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. Neuroreport 19 (11), 1163–1167.

Talmadge, C.L., Tubis, A., Long, G.R., Murphy, W.J., Tubis, A., 1990. Quantitative evaluation of limit-cycle oscillator models of spontaneous otoacoustic emissions. In Dallos, P., Geisler, C.D., Matthews, J. W., Ruggero, M., Steele, C.R., (Eds.), Mechanics and Biophysics of Hearing, Springer Verlag, New York, 235-242.

Talmadge, C.L., Long, G.R., Murphy, W.J., Tubis, A., 1993. New off-line method for detecting spontaneous otoacoustic emissions in human subjects. Hear. Res. 71, 170-182.

Talmadge, C.L., Tubis, A., Long, G.R., Piskorski, P., 1998. Modeling otoacoustic emission and hearing threshold fine structures. J. Acoust. Soc. Am. 104 (3), 1517-1543.

Tan, X., Beurg, M., Hackney, C., Mahendrasingam, S., Fettiplace, R., 2013. Electrical tuning and transduction in short hair cells of the chicken auditory papilla. J. Neurophysiol. 109, 2007-2020.

Taschenberger, G., Manley, G.A., 1997. Spontaneous otoacoustic emissions in the barn owl. Hear. Res. 110, 61-76.

Thomas, B., 1975. Microstructure of the pure-tone threshold. J. Acoust. Soc. Am 57, 26–27.

Tilney, M.S., Tilney, L.G., DeRosier, D.J., 1987. The distribution of jair cell bundle length and orientation suggest an unexpected pattern of hair cell stimulation in the chick cochlea. Hear. Res. 25, 141-151.

van der Heijden, M., Vavakou, A., 2021. Rectifying and sluggish: Outer hair cells as regulators rather than amplifiers. Hear. Res. 423, 108367, 1-12.

van Dijk, P. and Manley, A.M., 2001. Distortion product otoacoustic emissions in the tree frog Hyla cinerea. Hear. Res. 153, 14-22.

van Dijk, P. and Wit, H.P., 1990a. Amplitude and frequency fluctuations of spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 88 (4), 1779-1793.

van Dijk, P. and Wit, H.P., 1990b. Synchronization of spontaneous otoacoustic emissions to a $2f_1$-f_2 distortion product. J. Acoust. Soc. Am. 88, 850-856.

van Dijk, P., and Wit, H.P., 1998a. Correlated amplitude fluctuations of spontaneous otoacoustic emissions. J. Acoust. Soc. Am. 104, 336-343.

van Dijk, P. and Wit, H.P., 1998b. Synchronization of cubic distortion spontaneous otoacoustic emissions. J. Acoust. Soc. Am., 104 (1). 591-594.

van Dijk, P., and Wit, H.P., 1988. Phase-lock of spontaneous oto-acoustic emissions to a cubic difference tone. In Duifhuis, H., Horst, J. W., Wit, H. P., (eds.), Basic Issues in Hearing, Academic, London, 101-105.

van Dijk, P., Maat, B., de Kleine, E., 2011. The effect of static ear canal pressure on human spontaneous otoacoustic emissions: spectral width as a measure of the intra-cochlear oscillation amplitude. J. Assoc. Res. Otolaryngol. 12, 13-28.

van Dijk, P., Manley, A.G., Gallo, L., Pavusa, A., Taschenberger, G., 1996. Statistical properties of spontaneous otoacoustic emission in one bird and three lizard species. J. Coust. Soci. Am. 100, 2220-2227.

Vihla, M., Kiviniemi, K., Salmelin, R., 2002. Auditory cortical activation in the Finnish and Swedish speaking Finns: a magnetoencephalographic study. Neurosci. Lett. 322, 141–144.

von Békésy, G., 1960. Experiments in hearing. Wever, E.G. (Trans.), McGraw-Hill; New York.

von Helmholtz, H., 1862. Die Lehre von den Tonempfindung als physiologische Grundlage für die Theorie der Musik. Vieweg und Teubner Verlag.

Wan, I.K.K., Wong, L.L.N., 2002. Tympanometric norms for Chinese young adults. Ear Hear. 23 (5), 416–421.

Ward, W.D., 1955. Tonal monaural diplacusis. J. Acoust. Soc. Am. 27, 365-372.

Wegel, R.L. and Lane, C.E., 1924. The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. Phys. Rev. 23, 266-285.

Weisz, C.J.C., Lehar, M., Hiel, H., Glowatzki, E., Fuchs, P.A., 2012. Synaptic transfer from outer hair cells to type II afferent fibers in the rat cochlea. J. Neurosc. 32 (28), 9528-9536.

Whitehead, M.L., Kamal, N., Lonsbury-Martin, B.L., Martin, G.K., 1993. Spontaneous otoacoustic emissions in different racial groups. Scand. Audiol. 22, 3–10.

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis (Use R!), 2nd ed., Springer International Publishing, Cham.

Wilson, J.P., 1980. Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus. Hear. Res. 2, 233–252.

Wilson, J.P. and Sutton, G.J., 1981. Acoustic correlates of tonal tinnitus. In: Ciba Foundation Symposium 85-Tinnitus. Pitman Books Ltd, London, vol. 85, pp. 82–107.

Wong, P.C.M., Parsons, L.M., Martinez, M., Diehl, R.I., 2004. The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. J. Neurosci. 24 (41), 9153–9160.

Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat. Neurosci. 10, 420–422.

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models: Estimation of semiparametric generalized linear models. J. R. Stat. Soc. B (Stat. Methodol.) 73 (1), 3–36.

Wu, Y.C., Art, J.J., Goodman, M.B., Fettiplace, R., 1995. A kinetic description of the calcium-activated potassium channel and its application to electrical tuning of hair cells. Prog. Biophys. Mol. Biol. 63, 131-158.

Xia, A., Liu, X., Raphael, P.D., Applegate, B.E., Oghalai, J.S., 2016. Hair cell force generation does not amplify or tune vibrations within the chicken basilar papilla. Nat. Commun. 7 (13133), 1-12.

Zatorre, R.J., 2003. Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nat. Rev. Neurosci. 6 (7), 692–695.

Zatorre, R.J., Evans, A.C., Meyer, E., 1994. Neural mechanism underlying melodic perception and memory for pitch. J. Neurosci. Res. 14 (4), 1908–1919.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cogn. Sci. 6, 37–46.

Zettner, E.M. and Folsom, R.C., 2003. Transient emission suppression tuning curve attributes in relation to psychoacoustic threshold. J. Acoust. Soc. Am. 113, 2031–2041.

Zheng, J., Shen, W., He, D.Z.Z., Long, K.B., Madison, L.D., Dallos, P., 2000. Prestin is the motor protein of cochlear outer hair cells. Nature 405, 149-155.

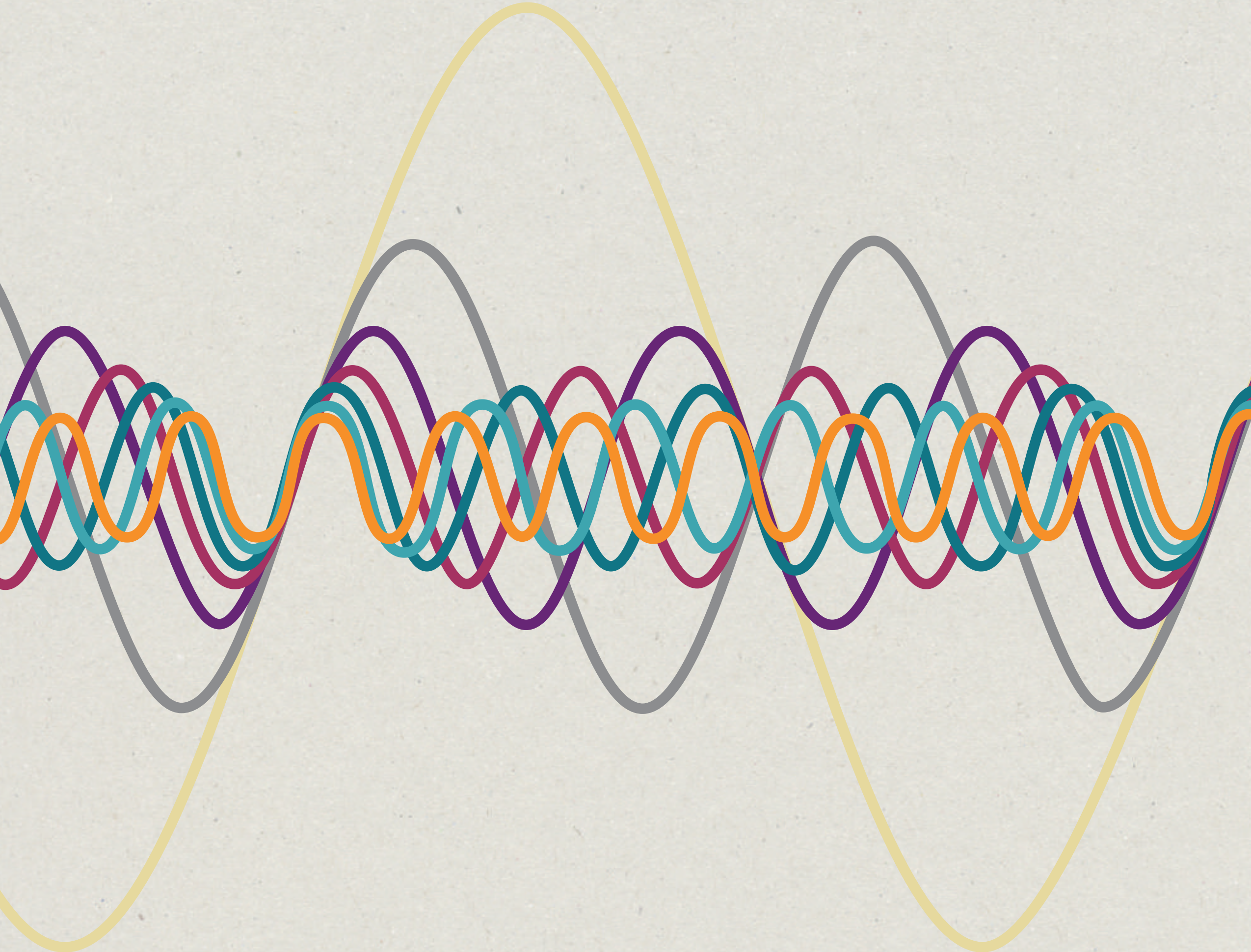
Zizz, C.A. and Glattke, T.J., 1988. Reliability of spontaneous optoacoustic emission sup- pression tuning curve measures. J. Speech Hear. Res. 31, 616-619.

Zurek, P.M., 1981. Spontaneous narrowband acoustic signals emitted by human ears. J. Acoust. Soc. Am. 69, 514–523.

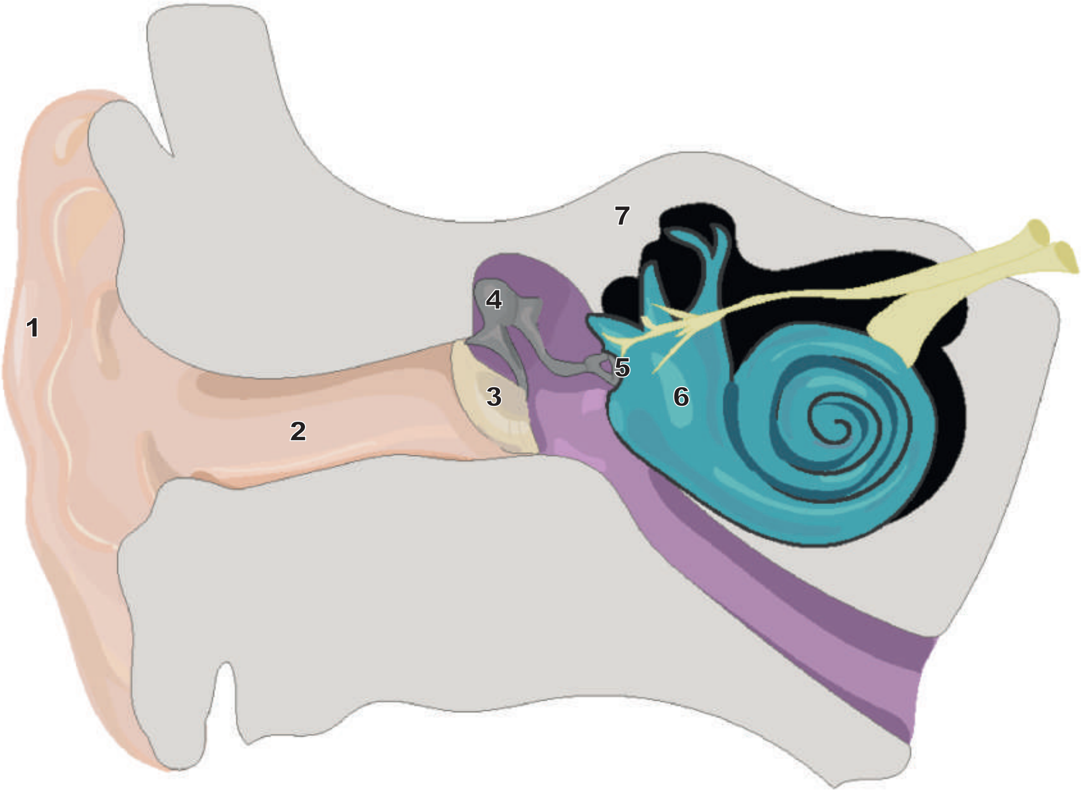
Zwicker, E., 1979. A model describing nonlinearities in hearing by active processes with saturation at 40 dB. Bioi Cybern 35-34, 243-250.

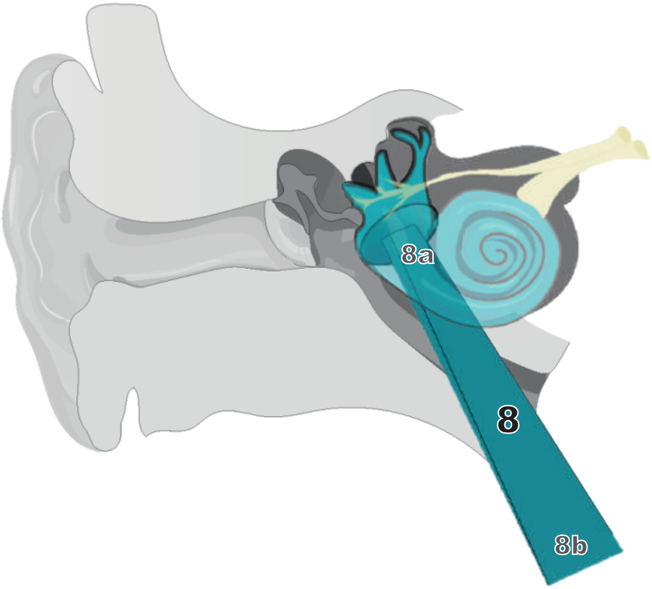
Zwicker, E., 1981. Dependence of level and phase of the ($2f_1$-f_2)-cancellation tone on frequency range, frequency difference, level of primaries, and subject. J. Acoust. Soc. Am. 70, 1277-1288.

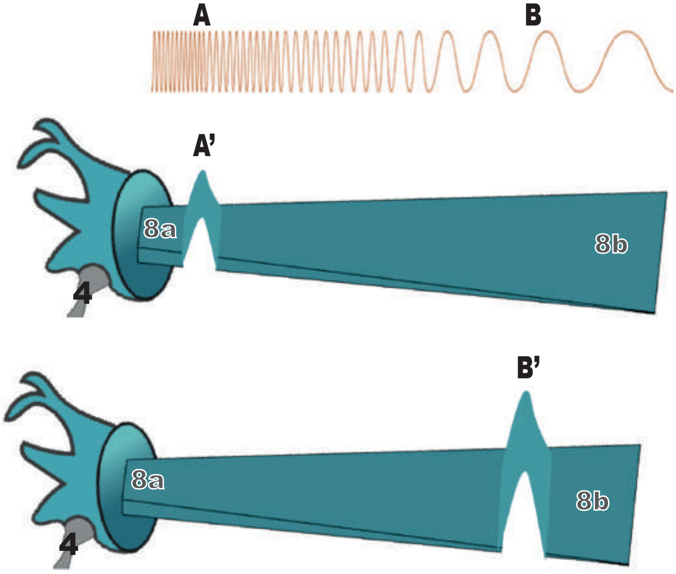
Zwicker, E. and Schloth, E., 1984. Interrelation of different oto-acoustic emissions. J. Acoust. Soc. Am. 75, 1148–1154.

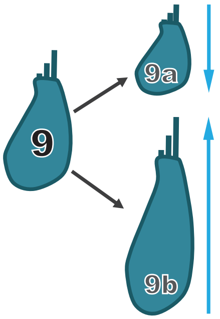


Summary
Samenvatting
Zusammenfassung


7
4
1
5
3
6
2


8a
8
8b


A
B
A’
8a
8b
4
B’
8a
8b
4


9
9a
9b

Summary

The goal of this summary is to explain the chapters of this thesis very illustratively. I am aiming to bring this very specific research field onto the radar of people who may be experienced in other fields of science or may not work in scientific research at all. I really hope to ‘build the bridge’ for layman-readers, and contribute to further understanding of this research and getting readers excited about it. Note that comparisons are made for illustrative reasons and have neither biological nor physical accuracy.

The way of sound through the human ear

Any sound that we naturally perceive consists of air pressure fluctuations. These fluctuations reach the pinna [1] and travel through the ear canal [2] to the tympanic membrane [3]. According to the presented sound pressure waves, the tympanic membrane is set into vibration. The tympanic membrane is in contact with the smallest bones of our body, the ossicles [4]. These three little bones form the ossicular chain that can move in a hammer-like manner. The ossicles are also in contact with the oval window [5], which represents the entrance to the inner ear. The inner ear contains a snail shaped structure [6] that curls almost three times around its own axis. This spiraled structure is called the cochlea, deriving from the Greek word for ‘snail shell’. The cochlea is a fluid filled bony structure, located within the skull [7]. When uncoiling the cochlea [8], the organ can be imagined as a tube that contains the basilar membrane. This membrane can be simplified as an array of structures with slightly different properties that merge into each other. The structural differences emerge from the base [8a] of the cochlea to its apex [8b], becoming wider and floppier.

The cochlea – the inner ear

To make the anatomy of the basilar membrane a bit more illustrative you may think of it as an uncoiled elastic bridge. Each place on the bridge has slight structural differences compared to the neighboring place. The structural differences of the bridge can be imaged as panels that smoothly blend into the neighboring one. The bridge is narrow and stiff at the beginning [8a] and becomes broader and floppier towards the other end [8b]. Thus, the structure of the bridge changes smoothly along its longitudinal axis.

Every part of the structure can be set into a maximal swinging motion. How a tone (presentation of one frequency) would set the membrane into motion can be illustratively mimicked by a person jumping on the bridge. When jumping on this bridge in a specific rhythm (*jumping-frequency*) that matches the resonance frequency of the local bridge structure, this construction would start to oscillate (swing). Since the bridge consists of many panels with slightly different structural properties, each part of the bridge will require a specific *jumping-frequency* (characteristic frequency) to reach its maximum oscillation. Of course, neighboring parts of this coupled structure will also move, but with smaller swinging levels (amplitude). Besides the exact *jumping frequency*, a sufficient force is also required to cause a maximum movement of the panel. Thus, in this model, the jumper would need to convince more people to jump at the same spot in synchrony.

(Of course, real bridges also have a resonance frequency. In fact, marching in lock step on a bridge is prohibited by law in Germany (§ 27 (6*) StVO), as the swinging can cause the bridge to collapse. Thus, for obvious reasons, jumping with many people on a bridge and trying to hit the resonance frequency is actually something you and your fellows should really not consider doing.)*

Back to our imaginary model: Since many companions jump with a specific frequency [A], a well-defined part of the bridge will oscillate [A']. When another team tries to set a different part of the bridge into motion [B'], their jumping rhythm consequently has to be different from the others [B].

Similar to the bridge, different areas of the membrane move according to the tone presented. Thus, when listening to a tone, a specific area of the membrane will oscillate. Therefore, when two tones of different frequency are presented ([A and B]; similar to the two teams of jumpers on a bridge from above), the membrane oscillates at two different regions [A' and B']. This representation of tones at a specific place is called *tonotopy*. Tonotopy can be understood as a map that indicates the location of vibration for any given frequency. Tones of high frequencies [A] cause basilar membrane movement at the base [A'], while the membrane oscillates at apical regions [B'] when lower frequency tones [B] are presented.

Amplification of sound within the cochlea

The membrane is located within the fluid filled system of the cochlea, not in air. Thus, the bridge in the analogy needs to be placed under water. When now trying to set a specific part of the drowned bridge into motion, on the one hand it is desired to have appropriate swinging levels and on the other hand to not set the entire bridge into vibration but the local part the jumping team is at. Hence, the jumping force needs local amplification.

Imagine springs [9] that are each area of the bridge. Each spring is also fixed to a roof-like structure above the bridge. When the team jumps, the spring shortens [9a] and pulls this specific part of the bridge upwards. When the team lands, however, the spring elongates [9b] and pushes the bridge downwards. The force applied by the spring matches the j*umping frequency* and therefore the peaks and valleys of the bridge oscillation. Moreover, the active force of the spring (shortening and elongation) amplifies the swinging motion of the bridge locally.

Of course, the membrane has no such springs, but specific cells that apply amplification. These cells are located at the surface of the membrane. Since they align in rows at the outer turn of the cochlea, they are termed outer hair cells [9]. These hair cells selectively sense the movement of the membrane and amplify it. These types of cells are able to modify their length, by shortening [9a] and stretching [9b] their cell bodies actively. Thus, hair cells generate a narrow oscillation peak on the membrane. Only with the amplification by the outer hair cells the human auditory system is sharply tuned to frequencies.

Spontaneous otoacoustic emissions

Going back to our bridge analogy, where the *jumping frequency* of a team represents the oscillation of a tone on the basilar membrane. Imagine that some springs would shorten and elongate spontaneously, without any *jumping force* applied. Indeed, those bridge regions would consequently oscillate even though nobody is jumping on them.

In the cochlea, outer hair cells can also oscillate spontaneously, resulting in membrane movements. Thus, while no tone is presented, areas of the membrane can oscillate. These oscillations are termed spontaneous otoacoustic emissions (SOAEs). You can picture that as the ear working in reverse.

Since the membrane is set into oscillation, the surrounding fluid within the cochlea moves, too. This motion propagates now through the cochlea also reaching the oval window [5] with the ossicles attached [4]. The motion transmitted through the bony chain, reaching the tympanic membrane [3]. The tympanic membrane works now as a speaker membrane, basically moving back and forth creating air pressure waves. When placing a miniature microphone into the ear canal [2], the pressure changes can be recorded as sound. Keep in mind that the tympanic membrane moves accordingly to the swinging that was originally generated by the outer hair cells. Basically, what we are recording are the SOAEs, a result of the spontaneous activity of the tiny hair cells. SOAEs are recorded as continuous pure tones. Consequently, when multiple SOAEs are present, multiple tones are recorded.

Note that the majority of normal hearing humans have SOAEs. People with SOAEs are usually not aware of them, as they are not perceived by the emitters themselves. They are also of such low intensity, that they are generally not heard by others. These characteristics of SOAEs make them distinct from the phenomenon of tinnitus. Tinnitus is usually a purely subjective perception and not recordable in the ear canal. Moreover, tinnitus is related to hearing loss, while SOAEs indicate the process of OHC amplification in healthy ears.

Application of spontaneous otoacoustic emissions

Remember that *tonotopy* refers to the membrane-map, in which each tone is assigned to a specific spatial place. When a tone is presented a specific area on the membrane is set into motion and thus a limited group of hair cells. This means that the ear reacts *frequency specific* to incoming sounds. This is necessary to hear out a single tone within a composition of multiple sounds, for example.

Since SOAEs are recorded as tones, each emission originates from activity at a specific narrow frequency region. But how can we measure how narrow this influenced region is? In terms of the bridge analogy, we would be interested in the expansion of the spontaneously active springs that push and pull a specific panel when generating the swinging. In other words: What is the *special range* of influence of this specific spring group?

To test that, the group of jumpers could start to jump next to the area of spontaneous oscillation. When the group jumps very close to the region of spontaneous oscillation, the emission will be disturbed by the *jumping frequency*. That means that the neighboring oscillation will start to vibrate with the same frequency that the group is applying. However, if only a few people and not the entire team is jumping, the applied force may not be sufficient to suppress the spontaneous oscillation next by. Especially, when the jumping team is further away from the spontaneously oscillating area, a stronger oscillation is then needed in order to interfere with the spontaneously active area. Thus, when the team is spatially far off, more jumping people are needed to generate a sufficient amplitude. When the spontaneously active region is spatially even further separated from the jumping team, the oscillation generated by the team may not reach the other area any more. Thus, the influence of the jumping team on the spontaneous active region is influenced by the spatial distance between the two oscillations and the amplitude generated by the jumping team.

The same holds for interaction of SOAEs and external tones on the membrane. When a tone is presented next to the SOAE, the level of the tone can be rather weak. With increasing frequency difference between the SOAE and the presented tone, the presented tone needs to be of higher level, to suppress the SOAE. Such measurements of SOAE suppression are used to test the *frequency selectivity* (tuning) at the place of the SOAE. Basically, this is testing the range of influence of the outer hair cells for this specific location on the basilar membrane. Note that this measurement does not involve any active participation of the tested person, it really measures the response of the ear to a given tone.

Study overview

SOAEs are present in all vertebrate classes, which allows us to evaluate the frequency selectivity of different species. The big advantage of SOAE suppression measurements is that there is no response of the tested individual needed. Thus, this method does not require training (which can also be quite time consuming) and is not as prone to fatigue or motivation dips as behavioral experiments that require active participation of the tested individual. The individual attitude towards the active test participation can influence the test results and is thus not as objective as SOAE suppression. Besides, the objectivity - the suppression of SOAEs is also a non-invasive measurement, the physical integrity is not affected.

In **Chapter 2** the frequency selectivity of the barn owl was measured, by suppressing SOAEs. Barn owls are the only birds we know so far that have SOAEs. Moreover, they do not suffer from age related hearing loss, as mammals do. The ability to regenerate damaged hair cells and the fact that their hearing range is similar to humans also put the barn owl into the focus of hearing research. Animal models allow measurements that are not feasible in humans, such as invasive neural recordings of the auditory nerve. Such neural recordings are of explicit value, because they allow the evaluation of information transfer from the ear to the brain, by measuring neural activity of single nerve fibers.

Moreover, animal models allow us to make comparisons between species, which is essential to understand different principles of the auditory system and their evolutionary development. The results of that study showed that barn owls neural tuning does not directly represent the frequency selectivity measured by SOAE suppression. Leading to the assumption that the results of the different methods of testing frequency selectivity cannot directly be compared.

Acoustical training is related to the improvement of auditory perception, as seen in professional musicians. Whether such training effects can also be measured objectively at the level of the cochlea was investigated in **Chapter 3**. Chinese is a tonal language in which changes in pitch may change the meaning of the word. The research question here was whether tonal language native speakers have enhanced frequency selectivity at cochlear level compared to non-tonal language native speakers. At the peripheral level of the ear no enhanced frequency selectivity was measured. Thus, possible auditory training effects by language acquisition are most likely to be based at central processing levels of the brain.

The influence of SOAEs on behavioral measurements was examined in **Chapter 4**. It was investigated whether frequency selectivity derived by the suppression of SOAEs represents the participants perception, since the spontaneous activity could have an influence on the frequency tuning. Participants with SOAEs did not have any different cochlear frequency selectivity compared to participants without SOAEs.

In **Chapter 5** the complex interactions of SOAEs with each other and with external tones were reviewed. Besides the SOAE suppression, external tones can for example also cause the emission to shift in frequency or generate further emissions. These interactions can also affect the perception of a tone by the participant. Consequently, such interactions need to be considered when measuring frequency tuning behaviorally as well.

Het gebruik van spontane otoakoestische emissies om frequentieselectiviteit te onderzoeken

Samenvatting

Het doel van deze samenvatting is om de hoofdstukken van dit proefschrift zeer illustratief uit te leggen. Ik streef ernaar om dit zeer specifieke onderzoeksgebied op de radar te krijgen van mensen die misschien ervaring hebben met andere wetenschapsgebieden of helemaal niet aan wetenschappelijk onderzoek werken. Deze samenvatting is geschreven voor een breed publiek om zo bij te dragen aan een beter begrip van dit onderzoek en om lezers enthousiast te maken. Sommige voorbeelden zijn gebruikt ter illustratie maar zijn biologisch of natuurkundig niet nauwkeurig.

De weg van geluid door het menselijk oor

Elk geluid dat we van nature waarnemen, bestaat uit luchtdrukgolven. Deze luchtdrukgolven bereiken de oorschelp [1] en gaan door de gehoorgang [2] naar het trommelvlies [3]. Door de geluidsdrukgolven wordt het trommelvlies in trilling gebracht. Dit membraan staat in contact met de kleinste botten van ons lichaam, de gehoorbeentjes [4]. Deze drie kleine botten vormen de gehoorbeentjesketen die op een hamerachtige manier beweegt als gevolg van de trilling van het trommelvlies. De gehoorbeentjes staan vervolgens in contact met het ovale venster [5]. Het ovale venster vormt de ingang van het binnenoor, een slakvormige structuur [6] die bijna drie keer om zijn eigen as krult. Deze spiraalvormige structuur wordt het slakkenhuis of cochlea genoemd, afgeleid van het Griekse woord voor ‘slakkenhuis’. Het slakkenhuis is een met vloeistof gevulde benige structuur die zich in de schedel [7] bevindt. Bij het denkbeeldig afrollen van het slakkenhuis kan het orgaan worden voorgesteld als een buis die het basilair membraan [8] bevat. Het basilair membraan kan worden vereenvoudigd als een reeks structuren met enigszins veranderende eigenschappen. De structurele veranderingen lopen van de basis [8a] van het slakkenhuis tot aan de top [8b], waarbij de structuren breder en slapper worden.

Het slakkenhuis – het binnenoor

Om de anatomie van het membraan wat te versimpelen kun je het zien als een afgerolde elastische brug. Elke plaats op de brug heeft kleine structurele verschillen met de naburige plaats. De structurele verschillen van de brug kunnen worden afgebeeld als panelen die vloeiend overgaan in de aangrenzende. De brug is in het begin smal en stijf [8a] en wordt naar het andere uiteinde breder en slapper [8b]. Zo verandert de structuur van de brug geleidelijk langs zijn lengteas.

Elk deel van de constructie kan in een maximale beweging worden gezet. Hoe een toon (bestaande uit één frequentie), het membraan in beweging zou zetten, kan illustratief worden nagebootst door een persoon die op de brug springt. Bij het springen op deze brug in een bepaald ritme (*springfrequentie*) dat overeenkomt met de resonantiefrequentie van de lokale brugstructuur, zou de constructie gaan oscilleren (trillen). Omdat de brug uit veel panelen bestaat met enigszins verschillende structurele eigenschappen, heeft elk gebied van de brug een specifieke *springfrequentie* (karakteristieke frequentie; [A, B]) nodig om zijn maximale trillingsuitslag te bereiken. Natuurlijk zullen aangrenzende delen van deze gekoppelde structuur ook bewegen, maar met kleinere uitslag (amplitude). Naast de exacte *springfrequentie* is er ook voldoende kracht nodig om een maximale beweging van het gebied te bewerkstelligen. In dit model zou de springer dus meer mensen moeten overtuigen om synchroon op dezelfde plek te springen.

(Natuurlijk hebben echte bruggen ook een resonantiefrequentie. In Duitsland is het daadwerkelijk verboden om op een brug te marcheren (§ 27 (6) StVO), omdat het slingeren de brug kan doen instorten. Vanzelfsprekend is met veel mensen op een brug springen en proberen de resonantiefrequentie te bereiken, niet iets is dat jij en je medemensen zouden moeten overwegen om te doen.)

Terug naar ons denkbeeldige model: wanneer veel metgezellen met een specifieke frequentie [A] springen, zal een goed gedefinieerd deel van de brug oscilleren [A']. Wanneer een ander team een ander deel van de brug in beweging probeert te brengen [B'], moet hun springritme dan ook anders zijn dan dat van de anderen [B].

Net als bij de brug bewegen verschillende delen van het basilair membraan afhankelijk van de frequentie van de aangeboden toon. Dus, bij het luisteren naar een toon, zal een specifiek gebied van het membraan oscilleren. Wanneer twee tonen van verschillende frequenties worden aangeboden [A en B], (vergelijkbaar met het voorbeeld van de twee teams springers op een brug), oscilleert het membraan in twee verschillende regio's [A' en B']. Deze weergave van tonen op een specifieke plaats wordt *tonotopie* genoemd. Tonotopie kan worden omschreven als een kaart die de locatie van trillingen voor een bepaalde frequentie aangeeft. Tonen van hoge frequenties [A] veroorzaken beweging aan het begin van het membraan [A'], terwijl het membraan oscilleert aan het eind [B'] wanneer tonen met een lagere frequentie worden aangeboden [B].

Versterking van geluid in het slakkenhuis

Het basilair membraan bevindt zich in het met vloeistof gevulde systeem van het slakkenhuis, en dus niet in lucht. De brug in de analogie moet dus onder water worden geplaatst. Wanneer men nu probeert om een heel specifiek deel van de ondergelopen brug in beweging te brengen is het belangrijk om ten eerste de geschikte amplitude te hebben en ten tweede om niet de hele brug in trilling te brengen maar alleen waar het lokale deel waar het springteam zich bevindt. Daarom heeft de springkracht lokale versterking nodig.

Stel je voor dat aan elk gebied van elk brugpaneel springveren [9] zijn bevestigd. Elke springveer is ook bevestigd aan een dakstructuur boven de brug. Wanneer het team springt, wordt de springveer korter [9a] en trekt dit het specifieke gebied van de brug omhoog. Wanneer het team landt wordt de springveer echter langer [9b] en duwt de brug naar beneden. De kracht die door de springveer wordt uitgeoefend komt overeen met de springfrequentie (spring en land) en dus met de pieken en dalen van de brug-oscillatie. Bovendien versterkt de actieve kracht van de springveer (verkorting en verlenging) de zwaaibeweging van de brug.

Het membraan heeft natuurlijk niet zulke springveren, maar wel specifieke cellen die versterking toepassen. Deze cellen bevinden zich aan het oppervlak van het membraan. Omdat ze in rijen zijn uitgelijnd aan de buitenste winding van het slakkenhuis worden ze buitenste haarcellen [9] genoemd. Deze haarcellen detecteren selectief de beweging van het membraan en versterken deze. Dit soort cellen kan hun lengte aanpassen door hun cellichamen actief in te korten [9a] en uit te rekken [9b]. Buitenste haarcellen zorgen dus voor een extra smalle oscillatiepiek op het membraan. Alleen met de versterking door de buitenste haarcellen is het menselijk gehoorsysteem scherp afgestemd op frequenties.

Spontane otoakoestische emissies

Teruggaand naar onze brug-analogie waar de *springfrequentie* van het team de trilling van de toon op het basilair membraan vertegenwoordigt. Stel je voor dat sommige springveren spontaan zouden inkorten en verlengen, zonder dat er enige *springkracht* wordt uitgeoefend. Inderdaad, die brugdelen zouden daardoor gaan oscilleren ook al springt er niemand op. Wanneer delen van de brug oscilleren wordt ook de omringende vloeistof in beweging gebracht.

In het slakkenhuis kunnen buitenste haarcellen ook spontaan oscilleren, wat resulteert in membraan-bewegingen. Dus, hoewel er geen toon wordt aangeboden, kunnen gebieden van het membraan oscilleren. Deze trillingen worden spontane otoakoestische emissies (SOAE's) genoemd. Je kunt je dat voorstellen als het oor dat omgekeerd werkt.

Omdat het membraan in trilling wordt gebracht beweegt ook de omringende vloeistof in het slakkenhuis. Deze beweging verspreidt zich nu door het slakkenhuis en bereikt ook het ovale venster [5] met de gehoorbeentjes [4] eraan vast. De beweging wordt nu overgebracht door de gehoorbeentjes en bereikt het trommelvlies [3]. Het trommelvlies werkt nu als een luidsprekermembraan dat heen en weer beweegt en luchtdrukgolven creëert. Bij het plaatsen van een miniatuurmicrofoon in de gehoorgang [2] kunnen de drukveranderingen als geluid worden waargenomen. Deze bewegingen van het trommelvlies worden nu alleen niet gegenereerd door geluid van buitenaf, maar door spontane oscillaties van de haarcellen in het binnenoor. Kortom, wat we opnemen zijn de SOAE's, een resultaat van de spontane activiteit van de kleine haarcellen. SOAE's worden geregistreerd als continue zuivere tonen. Als er dus meerdere SOAE's in een oor aanwezig zijn, worden er meerdere tonen opgenomen.

De meerderheid van de normaalhorende mensen heeft SOAE's. Echter, mensen met SOAE's zijn zich er meestal niet van bewust omdat ze niet door de "zenders" zelf worden waargenomen. Ze zijn ook van zo'n lage intensiteit dat ze over het algemeen niet door anderen worden gehoord. Deze kenmerken van SOAE's maken dat ze goed te onderscheiden zijn van het fenomeen tinnitus. Tinnitus is meestal een puur subjectieve waarneming en niet opneembaar in de gehoorgang. Bovendien is tinnitus gerelateerd aan gehoorverlies, terwijl SOAE's wijzen op het proces van haarcel-versterking in gezonde oren.

Toepassing van spontane otoakoestische emissies

Onthoud dat tonotopie verwijst naar de tonen-kaart van het membraan waar elke toon aan een specifiek gebied op het membraan wordt toegewezen. Wanneer een toon wordt aangeboden wordt een specifiek gebied op het membraan in beweging gebracht en dus een beperkte groep haarcellen. Dit betekent dat het oor *frequentiespecifiek* reageert op binnenkomende geluiden. Dit is nodig om bijvoorbeeld een enkele toon binnen een compositie van meerdere klanken te horen.

Aangezien SOAE's worden geregistreerd als tonen is elke emissie afkomstig van activiteit in een specifiek smal frequentiegebied. Maar hoe kunnen we meten hoe smal dit gebied is? In termen van de bruganalogie zouden we geïnteresseerd zijn in de uitrekking van de spontaan actieve springveren die een specifiek paneel duwen en trekken bij het genereren van het trillen. Met andere woorden: wat is het bereik van deze specifieke veergroep?

Om dat te testen, zou de groep springers kunnen gaan springen naast het gebied van spontane oscillatie. Als de groep heel dichtbij springt, wordt de spontane oscillatie verstoord door de springfrequentie. Dat betekent dat de naburige trilling zal gaan trillen met dezelfde frequentie die de groep toepast. Als er echter maar een paar mensen springen en niet het hele team, is de uitgeoefende kracht mogelijk niet voldoende om de spontane oscillatie te onderdrukken. Vooral wanneer het springteam ver weg is van het spontaan oscillerende gebied is een sterkere oscillatie nodig om het spontaan actieve gebied te verstoren. Dus, als het team op grotere afstand is, zijn er meer springende mensen nodig om een voldoende amplitude te genereren. Wanneer het spontaan actieve gebied op nog grotere afstand van het springteam is, kan de door het team gegenereerde trilling het andere gebied niet meer bereiken. De invloed van het springteam op het spontane actieve gebied wordt dus bepaald door de ruimtelijke afstand tussen de twee oscillaties en door de amplitude die door het springteam wordt gegenereerd.

Hetzelfde geldt voor de interactie van SOAE's en gepresenteerde tonen op het membraan. Wanneer een toon naast de SOAE wordt aangeboden kan een relatief zwakke toon de SOAE al onderdrukken. Met een toenemend frequentieverschil tussen de SOAE en de gepresenteerde toon moet de aangeboden toon van een hoger niveau zijn om de SOAE te onderdrukken. Dergelijke metingen van SOAE-onderdrukking worden gebruikt om de *frequentieselectiviteit* (*tuning*) op de plaats van de SOAE te meten. In principe is dit het testen van het bereik van de haarcellen voor deze specifieke locatie op het basilair membraan. Merk op dat deze meting geen actieve deelname van de geteste persoon vraagt, het meet echter alleen de reactie van het oor op een bepaalde toon.

Studieoverzicht

SOAE's zijn aanwezig in alle klassen van gewervelde dieren waardoor we de frequentieselectiviteit van verschillende soorten kunnen evalueren. Het grote voordeel van SOAE-onderdrukkingsmetingen is dat er geen actieve deelname van het individu nodig is. Deze methode vereist dus geen training (wat ook behoorlijk tijdrovend kan zijn) en is niet zo gevoelig voor vermoeidheid of afname in motivatie gedurende gedragsexperimenten. De individuele houding tijdens een actieve testdeelname kan de testresultaten beïnvloeden en is dus niet zo objectief als SOAE-onderdrukking. Daarnaast is de onderdrukking van SOAE's ook een niet-invasieve meting, de fysieke integriteit wordt niet aangetast.

In **Hoofdstuk 2** is de frequentieselectiviteit van de kerkuil gemeten door SOAE's te onderdrukken. Kerkuilen zijn de enige vogels die we tot nu toe kennen die SOAE's hebben. Bovendien lijden ze niet aan leeftijd-gerelateerd gehoorverlies zoals zoogdieren. Het vermogen om beschadigde haarcellen te regenereren en het feit dat hun gehoorbereik vergelijkbaar is met dat van mensen, plaatsen de kerkuil in de focus van gehoronderzoek.

Diermodellen maken metingen mogelijk die bij mensen niet haalbaar zijn zoals invasieve neurale metingen aan de gehoorzenuw. Dergelijke neurale opnames zijn van expliciete waarde, omdat ze de evaluatie van informatieoverdracht van het oor naar de hersenen mogelijk maken door de neurale activiteit van afzonderlijke zenuwvezels te meten. Bovendien stellen diermodellen ons in staat om vergelijkingen tussen diersoorten te maken wat essentieel is om de verschillende principes van het auditieve systeem en hun evolutionaire ontwikkeling te begrijpen. De resultaten van deze studie toonden aan dat de neurale afstemming van kerkuilen niet direct de frequentieselectiviteit weergeeft die wordt gemeten door SOAE-onderdrukking. Dit leidt tot de veronderstelling dat de resultaten van de verschillende methoden voor het testen van frequentieselectiviteit niet direct vergelijkbaar zijn.

ENG
NL
DE

Akoestische training leidt tot de verbetering van de auditieve waarneming, zoals dat wordt gezien bij professionele musici. Of dergelijke trainingseffecten ook objectief op het niveau van het slakkenhuis kunnen worden gemeten, is onderzocht in **Hoofdstuk 3**. Chinees is een tonale taal waarin veranderingen in toonhoogte de betekenis van het woord kunnen veranderen. De onderzoeksvraag hier was of sprekers met een tonale moedertaal een verbeterde frequentieselectiviteit hebben op cochleair niveau in vergelijking tot sprekers met een niet-tonale moedertaal. Op het perifere niveau van het oor werd geen verhoogde frequentieselectiviteit gemeten. Mogelijke effecten van auditieve training door taalverwerving zijn dus hoogstwaarschijnlijk gebaseerd op hogere verwerkingsniveaus in het brein.

De invloed van SOAE's op gedragsmetingen werd onderzocht in **Hoofdstuk 4**. Er werd onderzocht of frequentieselectiviteit zoals bepaald door de onderdrukking van SOAE's, overeenkomt met de perceptie van de deelnemers, aangezien de spontane activiteit invloed zou kunnen hebben op de frequentie-afstemming. Deelnemers met SOAE's hadden geen andere cochleaire frequentieselectiviteit in vergelijking met deelnemers zonder SOAE's.

In **Hoofdstuk 5** werden de complexe interacties van SOAE's met elkaar en met externe tonen besproken. Naast de SOAE-onderdrukking kunnen externe tonen er bijvoorbeeld ook voor zorgen dat een emissie in frequentie verschuift of dat er extra emissies gegenereerd worden. Deze interacties kunnen ook de perceptie van een toon door de deelnemer beïnvloeden. Daarom moet er met dergelijke interacties ook rekening worden gehouden bij het gedragsmatig meten van frequentieselectiviteit.

Die Anwendung von spontanen otoakustischen Emissionen zur Messung der Frequenselektivität

Zusammenfassung

Ziel dieser Zusammenfassung ist es, die Kapitel dieser Doktorarbeit anschaulich zu erläutern. Es ist mir ein besonderes Anliegen, dieses sehr spezifische Forschungsgebiet auch in den Blickpunkt derjenigen zu bringen, die entweder Erfahrung in anderen Wissenschaftsbereichen haben oder aber überhaupt nicht in der wissenschaftlichen Forschung tätig sind. Ich hoffe, für jene Laienleser *die Brücke bauen* zu können, um zu einem besseren Verständnis dieser Forschung beizutragen und vielleicht auch für dieses Themenfeld zu begeistern. Beachten Sie bitte, dass folglich Vergleiche zur Veranschaulichung vorgenommen werden, die mitunter weder biologisch noch physikalisch korrekt sind.

Der Weg des Schalls durch das menschliche Ohr

Jedes Geräusch, das wir wahrnehmen, besteht aus Luftdruckschwankungen (Schalldruckwellen). Diese Schwingungen erreichen die Ohrmuschel [1] und wandern durch den Gehörgang [2] zu einer Membran, dem Trommelfell [3]. Entsprechend den Druckschwankungen des Schalls wird das Trommelfell in Schwingung versetzt. Diese Membran steht in Verbindung mit den kleinsten Knochen unseres Körpers, den Gehörknöchelchen [4]. Die drei kleinen Knochen bilden die sogenannte Gehörknöchelchenkette, die sich hammerartig bewegen kann. Die Gehörknöchelchen sind im Kontakt mit dem ovalen Fenster [5], welches den Eingang zum Innenohr darstellt. Das Innenohr beinhaltet ein schneckenförmiges Gebilde [6], das sich fast dreimal um die eigene Achse windet. Diese spiralförmige Struktur wird Cochlea genannt, abgeleitet vom griechischen Wort für „Schneckenhaus“. Die Cochlea ist eine mit Flüssigkeit gefüllte knöcherne Struktur, die sich innerhalb des Schädels [7] befindet. Beim Abwickeln der Cochlea kann man sich das Organ als eine Röhre vorstellen, welche in ihrem Inneren die Basilarmembran [8] enthält. Diese Membran kann man sich vereinfacht als eine Anordnung von Strukturen mit jeweils leicht unterschiedlichen Eigenschaften vorstellen. Die strukturellen Unterschiede der Membran treten von der Basis [8a] bis zu ihrer Spitze [8b] auf, wobei die Membran zunehmend breiter und schlaffer wird.

Die Cochlea – das Innenohr

Um die Anatomie dieser Membran etwas anschaulicher zu machen, können Sie sich diese als eine abgewickelte elastische Brücke vorstellen. Jeder Bereich auf der Brücke weist gegenüber dem benachbarten Bereich leichte bauliche Unterschiede auf. Die strukturellen Unterschiede der Brücke kann man sich als Paneele mit unterschiedlicher Dicke und Festigkeit vorstellen, die jedoch fließend ineinander übergehen. Die Brücke ist am Anfang schmal und steif [8a] und wird zum Ende hin breiter und schlaffer [8b]. Somit ändert sich die Struktur der Brücke fließend entlang ihrer Längsachse [8].

Jeder Teil der Struktur kann in eine maximale Schwingbewegung versetzt werden. Wie ein Ton (*einzelne Frequenz*) die Membran in Bewegung setzen würde, kann am Beispiel einer Person verdeutlicht werden, die auf der Brücke springt. Beim Springen auf der Brücke in einem bestimmten Rhythmus (*Sprungfrequenz*), die der Schwingungsfrequenz (Resonanzfrequenz) des dortigen Brückenpaneels entspricht, würde die Konstruktion zu schwingen beginnen. Da die Brücke aus vielen Bereichen mit leicht unterschiedlichen strukturellen Eigenschaften besteht, benötigt jeder Paneelbereich eine bestimmte *Sprungfrequenz* (charakteristische Frequenz; [A, B]), um lokal die maximale Schwingung [A‘, B‘] zu erreichen. Natürlich werden sich auch benachbarte Bereiche dieser verbundenen Struktur bewegen, aber mit kleineren Schwingungen (Amplituden). Neben der genauen *Sprungfrequenz* ist auch eine ausreichende Kraft erforderlich, um eine maximale Bewegung des Bereiches zu bewirken. Bei diesem Modell müsste man demnach mehr Menschen davon überzeugen, synchron auf derselben Stelle zu springen.

(*Natürlich haben auch echte Brücken eine Resonanzfrequenz. Tatsächlich ist das Marschieren im Gleichschritt auf einer Brücke in Deutschland gesetzlich verboten (§ 27 (6) StVO), da das resultierende Schwingen die Brücke zum Einsturz bringen kann. Es gibt also offensichtliche Gründe, warum Sie nicht wirklich in Betracht ziehen sollten, mit vielen Mitmenschen auf einer Brücke zu springen und dabei zu versuchen, die Resonanzfrequenz dieser zu treffen.*)

Zurück zu unserem imaginären Modell: Da nun viele Personen in einer bestimmten Frequenz [A] springen, wird ein genau definierter Teil der Brücke schwingen [A‘]. Wenn ein anderes Team versucht, einen anderen Teil der Brücke in Bewegung zu setzen [B‘], muss dessen Sprungrhythmus folglich ein anderer sein [B].

Ähnlich wie bei der Brücke bewegen sich verschiedene Bereiche der Basilarmembran im Innenohr entsprechend dem dargebotenen Ton. Wenn Sie also einen Ton hören, schwingt ein bestimmter Bereich der Membran. Werden zwei Töne unterschiedlicher Frequenz dargeboten ([A und B], ähnlich wie die beiden Springteams auf einer Brücke von oben), schwingt die Membran in zwei verschiedenen Regionen [A‘ und B‘]. Diese Darstellung von Tönen an einem bestimmten Ort nennt man *Tonotopie*. Die Tonotopie kann als eine Karte auf der Membran verstanden werden, die den Ort der Schwingung für jede gegebene Frequenz anzeigt. Töne mit hohen Frequenzen [A] verursachen eine Bewegung der Membran an der Basis [A‘], während die Membran in Regionen an der Spitze schwingt [B‘], wenn Töne mit niedrigerer Frequenz dargeboten werden [B].

Schallverstärkung in der Cochlea

Die Basilarmembran befindet sich innerhalb des mit Flüssigkeit gefüllten Systems der Cochlea, ist also nicht von Luft umgeben. Daher muss die Brücke in unserem Beispiel unter Wasser getaucht werden. Wenn man nun versucht, einen bestimmten Teil der versenkten Brücke in Bewegung zu setzen, ist es zum einen erforderlich eine ausreichende Schwingung beizubehalten und zum zweiten nicht die gesamte Brücke in Schwingung zu versetzen, sondern nur den Bereich, an dem sich das Springteam befindet. Daher muss die Sprungkraft lokal verstärkt werden.

Stellen Sie sich Sprungfedern [9] vor, die oberhalb auf jedem Paneelbereich der Brücke angebracht sind. Jede Sprungfeder ist auch an einer dachähnlichen Struktur über der Brücke befestigt. Wenn das Team nun hochspringt, verkürzt sich die Sprungfeder [9a] in diesem Bereich und zieht diesen speziellen Teil der Brücke nach oben. Wenn das Team hingegen landet, verlängert sich die Feder [9b] und drückt die Brücke nach unten. Die von der Feder aufgebrachte Kraft entspricht der Sprungfrequenz (Hochspringen und Landen) des Teams und damit den Spitzen und Tälern der Brückenschwingung. Außerdem verstärkt die Wirkkraft der Feder (Verkürzung und Dehnung) lokal die Schwingbewegung der Brücke.

Natürlich hat die Basilarmembran des Innenohrs keine derartigen Sprungfedern, sondern spezielle Zellen, die eine Verstärkung des Tons gewährleisten. Diese Zellen befinden sich an der Oberfläche der Membran. Da sie an der äußeren Windung der Cochlea in Reihen angeordnet sind, werden sie als äußere Haarzellen [9] bezeichnet. Diese Haarzellen nehmen selektiv die Bewegung der Membran wahr und verstärken diese. Die Zellen sind dabei in der Lage, ihre Länge zu verändern, indem sie ihre Zellkörper aktiv verkürzen [9a] und strecken [9b]. Somit verstärken die äußeren Haarzellen die Schwingung auf der Membran und erzeugen einen schmalen Schwingungsbereich. Nur durch die Verstärkung dieser Zellen kann das menschliche Gehör so gut einzelne Frequenzen wahrnehmen.

Spontane otoakustische Emissionen

Zurück zu unserer Brücken-Analogie, wo die *Sprungfrequenz* des Teams die Schwingung eines Tons auf der Basilarmembran repräsentiert. Stellen Sie sich nun vor, dass sich einige Sprungfedern an einem Ort spontan verkürzen und verlängern, ohne dass jemand auf der Brücke springt. Folglich würde dieser Brückenbereich trotzdem schwingen.

In der Cochlea können die äußeren Haarzellen ebenfalls spontan oszillieren, was folglich zu Membranbewegungen führt. Somit können Bereiche der Membran schwingen, obwohl kein Ton dargeboten wird. Diese Schwingungen werden als *spontane otoakustische Emissionen* (SOAEs) bezeichnet. Sie können sich das so vorstellen, als würde das Ohr nun in entgegengesetzter Richtung arbeiten.

Da die Basilarmembran in Schwingung versetzt wird, bewegt sich auch die umgebende Flüssigkeit innerhalb der Cochlea. Diese Bewegung breitet sich durch die Cochlea aus und erreicht das ovale Fenster [5] mit den daran befestigten Gehörknöchelchen [4]. Die Knochenkette überträgt die Bewegung an das Trommelfell [3]. Das Trommelfell fungiert jetzt als ein Art Lautsprechermembran, die sich hin- und her bewegt und Luftdruckwellen erzeugt. Platziert man nun ein Miniaturmikrofon in den Gehörgang [2], können die Druckänderungen als Schall aufgezeichnet werden. Beachtlich ist dabei, dass sich das Trommelfell entsprechend der Schwingung bewegt, die ursprünglich von den Haarzellen erzeugt wurde. Was wir im Grunde aufzeichnen, sind die SOAEs, ein Ergebnis der spontanen Aktivität der winzigen Haarzellen des Innenohrs. SOAEs werden als kontinuierliche Einzelfrequenzen aufgezeichnet. Folglich werden mehrere Töne aufgezeichnet, wenn mehrere SOAEs vorhanden sind.

Die Mehrheit der normal hörenden Menschen hat SOAEs. Menschen mit SOAEs sind sich dieser normalerweise nicht bewusst, da die SOAEs von den Erzeugern selbst nicht wahrgenommen werden. Außerdem sind die Emissionen auch von so geringer Intensität, dass sie im Allgemeinen von anderen nicht gehört werden. Diese Eigenschaften von SOAEs unterscheiden sie vom Tinnitus-Phänomen. Tinnitus ist meist eine rein subjektive Wahrnehmung und nicht im Gehörgang detektierbar. Darüber hinaus steht Tinnitus im Zusammenhang mit Hörverlust, während SOAEs den Prozess der Haarzell-Verstärkung in gesunden Ohren verdeutlichen.

Anwendung spontaner otoakustischer Emissionen

Wir haben festgestellt, dass die Tonotopie eine Tonkarte der Membran ist, die jedem Ton einen bestimmten Bereich zuweist. Wenn also ein Ton präsentiert wird, wird ein bestimmter Bereich der Membran in Bewegung gesetzt und somit eine begrenzte Gruppe von Haarzellen. Das bedeutet, dass das Ohr *frequenzspezifisch* auf Töne reagiert. Dies ist beispielsweise notwendig, um innerhalb einer Komposition aus mehreren Klängen einen einzelnen Ton herauszuhören.

Da SOAEs als Töne aufgenommen werden, stammt jede Emission von der Aktivität in einem bestimmten schmalen Frequenzbereich auf der Membran. Aber wie können wir messen, wie eng dieser beeinflusste Bereich ist? In Bezug auf die Brückenanalogie interessiert uns die Ausdehnung der spontan wirkenden Sprungfedern, die bei der Erzeugung der Schwingung einen bestimmten Bereich ziehen oder drücken. Mit anderen Worten: Welchen lokalen Wirkungsbereich hat diese spezielle Federgruppe?

Um das zu testen, könnte das Team neben dem Bereich der spontanen Schwingung zu springen beginnen. Wenn die Gruppe sehr nah an dem Bereich der spontanen Schwingung springt, wird die spontane Schwingung durch die Sprungbewegung durch das Team gestört. Das bedeutet, dass die benachbarte spontane Schwingung der Emission mit der gleichen Frequenz zu schwingen beginnt, wie die Gruppe es in unmittelbarer Nähe vorgibt. Wenn jedoch nur wenige Personen und nicht das gesamte Team springen, reicht die aufgebrachte Kraft möglicherweise nicht aus, um die spontane Schwingung nebenan zu beeinflussen - insbesondere dann, wenn die Sprungmannschaft außerhalb des spontan schwingenden Bereichs, also auf einem weit entfernten Paneel, springt. Bei einer großen Entfernung ist dann eine stärkere Schwingung erforderlich, um den spontan aktiven Bereich zu beeinflussen. Wenn das Team also räumlich weit entfernt ist, werden mehr springende Personen benötigt, um eine ausreichende Schwingung zu erzeugen, als bei kurzer Distanz.

Wenn der spontan aktive Bereich räumlich noch weiter von der springenden Mannschaft entfernt ist, kann die von der Mannschaft erzeugte Schwingung den anderen Bereich nicht mehr erreichen, seien es auch noch so viele Teammitglieder. Der Einfluss der Sprungmannschaft auf den Bereich der spontanen Schwingung wird also durch den räumlichen Abstand der beiden Schwingungsbereiche und die von der Sprungmannschaft erzeugten Amplituden beeinflusst.

Dasselbe gilt für die Interaktion von SOAEs und präsentierten Tönen auf der Membran. Wenn ein Ton neben einer SOAE dargeboten wird, beide also eine ähnliche Frequenz haben, kann die Intensität des Tons schwach sein und er wird dennoch die Emission beeinflussen können. Mit zunehmenden Frequenzunterschied zwischen der SOAE und dem dargebotenen Ton muss der Ton eine höhere Intensität haben, um die Emission zu beeinflussen. Solche Messungen werden verwendet, um die Frequenzauflösung (*Frequenzselektivität*) am Ort der SOAE zu testen. Im Grunde wird hiermit der Einflussbereich der Haarzellen für diesen spezifischen Bereich auf der Basilarmembran getestet. Diese Messung bedarf keiner aktiven Teilnahme der getesteten Person, sondern misst die Reaktion des Ohrs auf einen bestimmten Ton.

Übersicht der durchgeführten Studien

SOAEs sind in allen Wirbeltierklassen vorhanden, was es uns ermöglicht, die Frequenzselektivität verschiedener Arten zu messen. Der große Vorteil von SOAE-Messungen besteht darin, dass keine Reaktion des Individuums notwendig ist. Somit erfordert diese Methode kein Training (das auch sehr zeitaufwändig sein kann) und ist nicht so anfällig für Ermüdung oder Motivationsverluste wie Verhaltensexperimente, die eine aktive Teilnahme der getesteten Person erfordern. Die individuelle Einstellung zur aktiven Testteilnahme kann die Testergebnisse beeinflussen und ist somit nicht so objektiv wie die oben beschriebene Methode. Zudem ist diese Methode zur Messung der Frequenzselektivität eine nicht-invasive Messung; die körperliche Unversehrtheit wird daher nicht beeinträchtigt.

In **Kapitel 2** wurde die Frequenzselektivität der Schleiereule gemessen, indem SOAEs durch präsentierte Töne unterdrückt wurden. Schleiereulen sind die einzigen bisher bekannten Vögel, welche SOAEs haben. Außerdem leiden sie nicht an altersbedingtem Hörverlust, wie es Säugetiere tun. Auch die Fähigkeit, geschädigte Haarzellen zu regenerieren und ihr, dem Menschen sehr ähnelnder, Hörbereich rücken die Schleiereule in den Fokus der Hörforschung.

Tiermodelle lassen Messungen zu, die beim Menschen nicht durchführbar sind, wie zum Beispiel invasive Nervenableitungen des Hörnervs, bei welchen die neuronale Aktivität einzelner Nervenfasern gemessen wird. Solche neuronalen Messungen sind von besonderem Wert, da sie die Erforschung der Informationsübertragung vom Ohr zum Gehirn ermöglichen. Darüber hinaus erlauben Tiermodelle Vergleiche zwischen Arten, was für das Verständnis verschiedener Prinzipien des Hörsystems und ihrer evolutionären Entwicklung unerlässlich ist. Die Ergebnisse der in Kapitel 2 durchgeführten Studie zeigten, dass bei Schleiereulen die neuronale Frequenzselektivität nicht direkt die Frequenzselektivität widerspiegelt, die durch SOAE-Unterdrückung gemessen wurde. Dies führt zu der Annahme, dass die Ergebnisse der verschiedenen Methoden zur Messung der Frequenzselektivität nicht direkt miteinander verglichen werden können.

Akustisches Training steht im Zusammenhang mit der Verbesserung der Hörwahrnehmung, wie sie bei professionellen Musikern beobachtet wird. Ob solche Trainingseffekte auch auf der Ebene der Cochlea objektiv messbar sind, wurde in **Kapitel 3** untersucht. Chinesisch ist eine tonale Sprache, bei der Tonhöhenänderungen die Bedeutung des Wortes verändern können. Die Forschungsfrage hier war, ob Muttersprachler*innen einer Tonsprache im Vergleich zu nicht-tonalen Muttersprachlern*innen eine verbesserte Frequenzselektivität auf der Cochlea-Ebene aufweisen. Auf der peripheren Ebene des Ohrs wurde keine erhöhte Frequenzselektivität gemessen. Mögliche Trainingseffekte durch den Spracherwerb ruhen also wahrscheinlich auf höheren (zentralen) Verarbeitungsebenen des Gehirns.

Der Einfluss von SOAEs auf akustische Verhaltensmessungen wurde in **Kapitel 4** untersucht. Es wurde getestet, ob die durch die Unterdrückung von SOAEs abgeleitete Frequenzselektivität des Innenohrs auch die eigentliche Wahrnehmung der Studienteilnehmenden widerspiegelt, da die spontane Aktivität einen Einfluss auf die Frequenzbestimmung haben könnte. Teilnehmende mit SOAEs hatten keine unterschiedliche Cochlea-Frequenzselektivität im Vergleich zu Teilnehmenden ohne SOAEs.

In **Kapitel 5** wurden die komplexen Wechselwirkungen von SOAEs untereinander und mit externen Tönen dargestellt. Neben der SOAE-Unterdrückung können beispielsweise auch externe Töne eine Frequenzverschiebung der Emission bewirken oder weitere Emissionen erzeugen. Diese Interaktionen können auch die Wahrnehmung eines Tons beeinflussen. Folglich müssen solche möglichen Wechselwirkungen auch bei der Messung berücksichtigt werden.

About the author

As a further course of education Sina Engler started to study Biology at the Carl von Ossietzky University in Oldenburg (Germany) in 2011. During this time, she performed an independent research project in Costa Rica, where she studied mother-pup recognition of bats. This project became a profound experience, leading her into the field of acoustic research.

In 2016 she completed her Master's degree, working in the neuroscience lab of prof. Christine Köppl. Enjoying the great academic and friendly atmosphere she decided to continue in scientific research and started a PhD project at the Ear, Nose, and Throat department of the University Medical Center Groningen (Netherlands), which resulted in the current dissertation. She was supervised by prof. Pim van Dijk, prof. Nomdo Jansonius and dr. Emile de Kleine. While finishing her thesis she moved back to Germany and started working in an engineering office, where she continues to work with bats.

Acknowledgements

During my time here in Groningen I was lucky to meet people and already knowing others that supported me in several ways. Being part of the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie-COFUND (EGRET) allowed me to become part of a comprehensive PhD program with many opportunities to build up extensive (academic) networks and gaining international working experience. I am very grateful for this experience.

Even though I wrote this thesis myself it's nothing I did on my own. What does that mean? Well, my personal goal of the past years was to develop through my time as a PhD student to scientist that critically questions the known and never disregards the own surrounding. I am thankful that my surrounding also never disregarded me and I am grateful for all the support I received. In the following I would like to point out some of the people that highly contributed to my success.

First of all, I thank you **Pim** for always being such an open minded, honest, and trustful supervisor and person. This relationship made it very easy to come out with difficulties and 'call the issue at its name'. I could always rely on your support and you gave me the freedom to develop my own scientific working behavior and personality. We often started meetings in Dutch and then switched to English, as *"German is just another small dialect of Dutch."* (Pim van Dijk, meeting 2020) - we simply neglected it. I will surely miss starting each meeting with a cup of coffee and a brief chat with you and Emile.

Emile, you have been my supervisor for the last years and I always enjoyed your fruitful thoughts, critical questioning and very empathic personality. I appreciate your calming character also when you had to be stressed finding yourself in a meeting in between the lab meeting and the ENT on-call duty. You gave me the opportunity to also learn more about the clinical work, thank you.

Dear **Nomdo** we interacted mainly during EGRET seminars and meetings, as our research fields did not have big topical intersections. You always provided very useful thoughts that assist a critical reflection of the own work. Thank you for your support.

I also thank the reading committee: **prof. J. C. M. Smits, prof. J. M. C. van Dijk,** and **prof. J. L. Verhey** for evaluating my dissertation.

Being a PhD in the ENT department meant to have many **PhD colleagues**. We grew together as a supportive community and had 'gezellige' lunch breaks. During the

COVID-19 crises we tried to continue with coffee breaks on skype. We met for ‘distance walks’ in and around Groningen. The relationship to all of you has always been warm and friendly. I have learned a lot from each one of you, no matter if it was that living as a vegan is apparently enjoyable or something else. I wish you all the best for your projects and success for your dissertations. Thank you all for making each UMCG coffee a good one.

A special thank you goes to **Dora**, we had very fruitful discussions and very nice talks. Over the years we built up a strong friendship and I am positive that at some point we will really cook a fabulous Croatian-German food intermezzo and not just dream about it while ordering a pizza.

Thank you - **Daniël** for all the nice conversations we had, whether they were scientific or not. I will remember our honest chats that smoothly graduated from one topic into another and wish you all the best.

I would like to emphasize that our lab-meetings have been very diverse in topics and it was always a good opportunity to hear critical thoughts and new ideas. I especially want to thank **Wiebe Horst** and **Diek Duifhuis** for their inspiring work and the impressive enthusiasm for science and young scientists. *Unfortunately, Diek passed away in July 2022.*

I am also thankful for the support within the entire **ENT department**, the clinical staff at the policlinic as well as our secretaries, that all keep the system running smoothly.

Paolo, we had enjoyable coffee breaks at the blue patio, which I will miss. You have been a great support and I will always remember: When you put an *if* never forget the *end*.” (#MatlabRocks)

Moreover, I would like to thank the **EGRET(+) community**, including PhD students, supervisors and coordinators. I experienced this time as a great opportunity to learn not only more about other fields of science, but also about other cultures and countries. We traveled a lot together and built up some interdisciplinary cooperation’s. Some of us became close friends, I am glad I meet every single one of you and wish you all the best for your future.

I also want to thank all the **participants** for their voluntary contribution and their interest in this research. Thank you all for making the projects possible.

Now it’s time to thank you **Bianca** for always being such a trustful person and making this wonderful friendship something I can always rely on. I am looking forward to many more memories we will share in the future. You are my best friend, and you always will be.

A very special thank you goes to **my family** who was never tired when I used one of the family gatherings to point out some amazing scientific facts. Thank you all for your kind and cheering words.

Thank you - **Monty**, for being the best doggy I can imagine. You have been a very important companion to zoom out the academic bubble and gather new energy for upcoming steps. You are such a good boy!

Liebe **Eltern** - es gibt keine Worte, die auch nur annähernd beschreiben könnten, wie dankbar ich euch für all eure bedingungslose Liebe und Hilfe während meines ganzen Lebens bin. Mein Weg zum Doktortitel war weder gerade noch eben, aber ihr beide habt stets sichergestellt, dass ich all den *Proviant* dabei habe, um mein Ziel zu erreichen. In diesem Sinne war mein Rucksack auf dieser Reise stets gefüllt mit Liebe, Fürsorge und Unterstützung von euch beiden. Danke für alles.