

**Temporal perception in the Mongolian gerbil: effects of  
aging and experimentally induced synaptopathy**

Von der Fakultät für Medizin und  
Gesundheitswissenschaften der Carl von Ossietzky  
Universität Oldenburg zur Erlangung des Grades und Titels  
eines

**Doktors der Naturwissenschaften (Dr. rer. Nat.)**

angenommene Dissertation  
**von Herrn Henning Oetjen**

geboren am 03.02.1988 in Bremen

Erstgutachter: Prof. Dr. Georg M. Klump

Weitere Gutachter:  
Prof. Dr. Peter Heil  
Prof. Dr. Christine Köppl

Tag der Disputation

16.07.2021



## Summary

The global number of people suffering from disabling hearing loss is estimated to double from 466 million to 900 million within the next 30 years. One aspect of hearing loss is temporal perception deficits. Temporal perception refers to the perception of variation in acoustic stimuli over time and is an important component for the perception of speech, pitch, loudness and sound localization. Animal and human studies suggested that temporal perception is negatively affected by general aging processes and synaptopathy, the dysfunction or loss of synapses between inner hair cells and auditory nerve fibers. Research in humans is mainly limited to non-invasive measures. Thus, it is necessary to rely on animal data in the same species. By relating electrophysiological and behavioral measures of sensitivity with direct histological measures, the consequences of synaptopathy and aging on temporal perception can be investigated. The overall aim of the present thesis was to conduct such a comparison by investigating temporal perception with focus on aging and experimentally induced synaptopathy in the Mongolian gerbil.

The goal of the first study was to investigate temporal perception and its relation to neural representation in young normal-hearing gerbils using Schroeder-phase complexes. These stimuli are harmonic tone complexes with a certain phase relation. Discrimination of Schroeder-phase complexes with different fundamental frequencies and duty cycles were examined. The behavioral discrimination of Schroeder-phase complexes was compared to neuronal single-unit responses of Schroeder-phase complexes in the auditory nerve and inferior colliculus. Two different experiments were conducted in the gerbil with either only acoustic temporal fine structure cues being available for discrimination or both acoustic temporal fine structure and envelope cues being available. The discrimination with only temporal fine structure cues being available was additionally investigated in young normal-hearing humans. This study provided the first comparison of behavioral and neural sensitivities for discriminating Schroeder-phase complexes within the same species. The main findings demonstrated gerbils and humans were able to discriminate behaviorally Schroeder-phase complexes for fundamental frequencies up to 200 Hz. In the case of only temporal fine structure cues being available for discrimination, the correlations between behavioral and neural temporal fine structure- and envelope-based sensitivities were high for both auditory

nerve and inferior colliculus responses. With temporal fine structure and envelope cues available for discrimination, the correlations between behavioral were highest for either rate- or envelope-based auditory nerve sensitivities, and for temporal fine structure- and envelope-based inferior colliculus sensitivities. These findings demonstrate temporal features of Schroeder-phase complexes are preserved to some extent in the gerbil's auditory pathway up to the inferior colliculus.

The goal of the second study was to investigate the effects of age on temporal perception when discriminating Schroeder-phase complexes based on temporal fine structure cues. The results of young normal-hearing gerbils of the first study were used as a benchmark. Auditory brainstem responses to click stimuli were recorded to verify the hearing status of each individual. The results demonstrate that old gerbils can discriminate Schroeder-phase complexes for fundamental frequencies up to 200 Hz equal to young gerbils. For fundamental frequencies below 200 Hz, the old gerbils showed age-related decreased sensitivity in temporal perception. The auditory brainstem responses revealed age-related reduced wave I amplitudes, indicating synaptopathy in the old gerbils. The observed decrease in behavioral sensitivity with increasing fundamental frequency suggests, however, that age-related deficits in central coding may also be responsible for a decline in temporal perception, possibly caused by an altered inhibitory network in the central auditory system.

In the third study, the effect of experimentally induced synaptopathy and aging on temporal perception was investigated behaviorally using the TFS1 test relying on discrimination of harmonic and frequency-shifted inharmonic tone complexes. The effect of age was examined by comparing the sensitivity of young and old gerbils. Additionally, young gerbils, treated with ouabain to degenerate the low-spontaneous rate fibers were used for investigating the effect of synaptopathy. The hearing status in the gerbils was verified with auditory brainstem responses to click stimuli. The results showed that the auditory brainstem response amplitudes were decreased for all groups with respect to the young untreated gerbils, indicating synaptopathy due to aging and ouabain treatment. In line with this, the temporal perception was negatively affected by age. However, no differences were observed within the groups of young gerbils including the ouabain treated groups, suggesting low-spontaneous rate fiber loss did not affect temporal perception. Overall, the results suggest that temporal perception

deficits in the TFS1 test are due to central processing deficits, possibly resulting from altered central inhibition.

This thesis highlights the benefit of animal studies in auditory research. It stresses the need for future research on the effects of aging on temporal perception with a focus on the central auditory system.

## **Zusammenfassung**

Es wird geschätzt, dass sich die Zahl der Menschen, die weltweit an einer Schwerhörigkeit leiden, innerhalb der nächsten 30 Jahre von 466 Millionen auf 900 Millionen verdoppeln wird. Ein Aspekt der Schwerhörigkeit sind Defizite in der zeitlichen Wahrnehmung. Die zeitliche Wahrnehmung bezieht sich auf die Wahrnehmung von Variationen akustischer Reize über die Zeit und ist eine wichtige Komponente für die Wahrnehmung von Sprache, Tonhöhe, Lautheit und Schalllokalisierung. Tier- und Humanstudien deuten darauf hin, dass die zeitliche Wahrnehmung durch allgemeine Alterungsprozesse und Synaptopathie, die Dysfunktion oder den Verlust von Synapsen zwischen inneren Haarzellen und Hörnervenfasern, negativ beeinflusst wird. Die Forschung am Menschen ist hauptsächlich auf nicht-invasive Messungen beschränkt. Daher ist es notwendig, sich auf Tierdaten der gleichen Spezies zu stützen. Indem elektrophysiologische und verhaltensbiologische Messungen der Sensitivität mit direkten histologischen Messungen in Beziehung gesetzt werden, können die Konsequenzen von Synaptopathie und Alterung auf die zeitliche Wahrnehmung untersucht werden. Das übergeordnete Ziel der vorliegenden Arbeit war es, einen solchen Vergleich durchzuführen, indem die zeitliche Wahrnehmung mit Fokus auf Alterung und experimentell induzierter Synaptopathie bei der Mongolischen Wüstenrennmaus untersucht wurde.

Das Ziel der ersten Studie war es, die zeitliche Wahrnehmung und ihre Beziehung zur neuronalen Repräsentation bei jungen normalhörenden Wüstenrennmäusen unter Verwendung von Schroeder-Phasen-Komplexen zu untersuchen. Diese Stimuli sind harmonische Tonkomplexe mit einer bestimmten Phasenrelation. Untersucht wurde die Diskriminierung von Schroeder-Phasen-Komplexen mit unterschiedlichen Grundfrequenzen und Aktivierungsdauern. Die Verhaltensunterscheidung von Schroeder-Phasen-Komplexen wurde mit neuronalen Einzelantworten von Schroeder-Phasen-Komplexen im Hörnerv und im Colliculus inferior verglichen. Zwei verschiedene Experimente wurden in der Wüstenrennmaus durchgeführt, wobei entweder nur akustische zeitliche Feinstruktur-Informationen für die Diskriminierung zur Verfügung standen oder sowohl akustische zeitliche Feinstruktur- als auch Einhüllenden-Informationen zur Verfügung standen. Die Diskriminierung, bei der nur zeitliche Feinstruktur-Informationen zur Verfügung stehen, wurde zusätzlich bei

jungen normalhörenden Menschen untersucht. Diese Studie lieferte den ersten Vergleich von Verhaltens- und neuronalen Empfindlichkeiten für die Unterscheidung von Schroeder-Phasen-Komplexen innerhalb derselben Spezies. Die Hauptergebnisse zeigten, dass Wüstenrennmäuse und Menschen in der Lage waren, Schroeder-Phasen-Komplexe für Grundfrequenzen bis zu 200 Hz im Verhalten zu unterscheiden. Wenn nur zeitliche Feinstruktur-Informationen für die Unterscheidung zur Verfügung standen, waren die Korrelationen zwischen verhaltensbezogenen und neuronalen zeitlichen Feinstruktur- und Einhüllenden-basierten Empfindlichkeiten sowohl für Hörnerven- als auch für Colliculus inferior-Antworten hoch. Wenn zeitliche Feinstruktur- und Einhüllenden-Informationen für die Unterscheidung zur Verfügung standen, waren die Korrelationen zwischen dem Verhalten am höchsten für entweder raten- oder einhüllendenbasierte Empfindlichkeiten des Hörnervs und für zeitliche Feinstruktur- und einhüllendenbasierte Empfindlichkeiten des Colliculus inferior. Diese Ergebnisse zeigen, dass temporale Merkmale der Schroeder-Phasen-Komplexe in der Hörbahn der Wüstenrennmaus bis zu einem gewissen Grad bis zum Colliculus inferior erhalten sind.

Das Ziel der zweiten Studie war es, die Auswirkungen des Alters auf die zeitliche Wahrnehmung bei der Unterscheidung von Schroeder-Phasen-Komplexen auf der Basis von zeitlicher Feinstruktur-Informationen zu untersuchen. Die Ergebnisse junger normalhörender Wüstenrennmäuse aus der ersten Studie wurden als Benchmark verwendet. Auditive Hirnstammantworten auf Klickreize wurden aufgezeichnet, um den Hörstatus jedes Individuums zu verifizieren. Die Ergebnisse zeigen, dass alte Wüstenrennmäuse Schroeder-Phasen-Komplexe für Grundfrequenzen bis 200 Hz genauso diskriminieren können wie junge Rennmäuse. Für Grundfrequenzen unter 200 Hz zeigten die alten Wüstenrennmäuse eine altersbedingt verminderte Empfindlichkeit in der zeitlichen Wahrnehmung. Die auditorischen Hirnstammantworten zeigten altersbedingt reduzierte Welle I Amplituden, was auf eine Synaptopathie bei den alten Wüstenrennmäusen hinweist. Die beobachtete Abnahme der Verhaltensempfindlichkeit mit zunehmender Grundfrequenz deutet jedoch darauf hin, dass altersbedingte Defizite in der zentralen Kodierung ebenfalls für eine Abnahme der zeitlichen Wahrnehmung verantwortlich sein könnten, möglicherweise



verursacht durch ein verändertes inhibitorisches Netzwerk im zentralen auditorischen System.

In der dritten Studie wurde die Auswirkung der experimentell induzierten Synaptopathie und des Alterns auf die zeitliche Wahrnehmung mit Hilfe des TFS1 Tests, der auf der Unterscheidung von harmonischen und frequenzverschobenen unharmonischen Tonkomplexen beruht, verhaltensbiologisch untersucht. Der Effekt des Alters wurde durch den Vergleich der Sensitivität von jungen und alten Wüstenrennmäusen untersucht. Zusätzlich wurden junge Wüstenrennmäuse, die mit Ouabain behandelt wurden, um die Fasern mit niedriger Spontanrate zu degenerieren, verwendet, um den Effekt der Synaptopathie zu untersuchen. Der Hörstatus wurde bei den Wüstenrennmäusen anhand der Reaktionen des auditorischen Hirnstamms auf Klickreize überprüft. Die Ergebnisse zeigten, dass die Amplituden der auditorischen Hirnstammantworten bei allen Gruppen im Vergleich zu den jungen, unbehandelten Wüstenrennmäusen vermindert waren, was auf eine Synaptopathie aufgrund der Alterung und der Behandlung mit Ouabain hinweist. Im Einklang damit wurde die zeitliche Wahrnehmung durch das Alter negativ beeinflusst. Es wurden jedoch keine Unterschiede innerhalb der Gruppen junger Rennmäuse einschließlich der mit Ouabain behandelten Gruppen beobachtet, was darauf hindeutet, dass der Verlust von Fasern mit niedriger Spontanrate die zeitliche Wahrnehmung nicht beeinflusst. Insgesamt legen die Ergebnisse nahe, dass zeitliche Wahrnehmungsdefizite im TFS1 Test auf zentrale Verarbeitungsdefizite zurückzuführen sind, die möglicherweise aus einer veränderten zentralen Inhibition resultieren.

Diese Arbeit unterstreicht den Nutzen von Tierstudien in der auditorischen Forschung. Sie unterstreicht die Notwendigkeit zukünftiger Forschung zu den Auswirkungen des Alterns auf die zeitliche Wahrnehmung mit einem Fokus auf das zentrale auditorische System.

# Contents

Summary.....	III
Zusammenfassung.....	VI
Abbreviations.....	1
<b>1 Introduction .....</b>	<b>2</b>
1.1 Physiological representation of temporal coding .....	3
1.2 Effects of age-related and noise-induced hearing loss on temporal perception ..	8
1.3 The Mongolian gerbil as an animal model for the experimental evaluation of temporal perception .....	13
1.4 Thesis outline and aim.....	14
<b>2 Chapter Schroeder-phase discrimination in young gerbils.....</b>	<b>16</b>
2.1 Abstract.....	17
2.2 Introduction.....	18
2.3 Materials and methods .....	20
2.3.1 Subjects.....	20
2.3.2 Schroeder-phase-complex stimuli .....	21
2.3.3 Auditory-nerve recordings.....	23
2.3.4 Inferior Colliculus recordings.....	25
2.3.5 Neural recordings - Data Analysis.....	27
2.3.6 Behavioral Experiments .....	30
2.3.6.1 Apparatus and stimulus generation: gerbils .....	30
2.3.6.2 Behavioral procedure: gerbils .....	30
2.3.6.3 Behavioral procedure: humans.....	32
2.3.6.4 Behavioral data analysis.....	32
2.3.6.5 Neural and behavioral statistical analysis.....	32
2.4 Results .....	33
2.4.1 Neural responses to sweeps that differed in direction.....	33
2.4.2 Discrimination of sweep direction based on AN responses .....	37

2.4.3	Discrimination of sweep direction based on IC responses .....	39
2.4.4	Behavioral discrimination of sweep direction by gerbils and humans.....	40
2.4.5	Discrimination of sweep velocity based on AN responses .....	40
2.4.6	Discrimination of sweep velocity based on IC responses .....	46
2.4.7	Discrimination of sweep velocity based on behavioral responses by gerbils .....	48
2.4.8	Correlations between behavioral and neural discrimination.....	48
2.4.8.1	Sweep-direction experiment.....	48
2.4.8.2	Sweep-velocity experiment.....	51
2.5	Discussion.....	54
2.5.1	Mechanisms underlying neural discrimination .....	54
2.5.2	Relating single-unit responses to behavioral performance.....	56
2.5.3	Comparative analysis of perception of SCHR complex stimuli .....	57
<b>3</b>	<b>Chapter Schroeder-phase discrimination in old gerbils .....</b>	<b>58</b>
3.1	Abstract.....	59
3.2	Introduction.....	60
3.3	Materials and methods .....	62
3.3.1	Subjects.....	62
3.3.2	Auditory brainstem response measurement.....	62
3.3.3	Behavioral apparatus and stimulus generation.....	63
3.3.4	Procedure .....	65
3.3.5	Behavioral data Analysis.....	67
3.4	Results .....	67
3.4.1	Auditory brainstem responses .....	67
3.4.2	Behavioral sensitivity.....	68
3.5	Discussion.....	71
3.5.1	Temporal perception with age .....	71
3.5.2	Physiological constraints of temporal perception.....	72
<b>4</b>	<b>Chapter TFS1 test: a comparison of Mongolian gerbils with normal and compromised cochlear function.....</b>	<b>75</b>

4.1	Abstract .....	76
4.2	Introduction.....	77
4.3	Material and Methods .....	79
4.3.1	Subjects.....	79
4.3.2	Ouabain treatment .....	80
4.3.3	Auditory brainstem response recording.....	80
4.3.4	Behavioral apparatus and stimuli.....	82
4.3.5	Procedure .....	86
4.3.6	Statistical Analysis .....	88
4.4	Results .....	88
4.4.1	Auditory brainstem response .....	88
4.4.2	Behavior sensitivity.....	91
4.5	Discussion.....	95
4.5.1	TFS1 sensitivity in young and old subjects: comparison between humans and gerbils .....	95
4.5.2	Effects of synaptopathy on temporal perception.....	98
4.5.3	Effects of central processing deficits on temporal perception.....	100
<b>5</b>	<b>General Discussion .....</b>	<b>102</b>
5.1	Summary of the experimental chapters.....	103
5.2	Common findings from the three experiments.....	105
5.3	Comparison of temporal perception in gerbils and humans.....	106
5.3.1	SCHR complex discrimination in old humans .....	106
5.3.2	Loss of low-SR fibers with aging in humans.....	107
5.3.3	Effect of low-SR fibers loss on temporal perception in humans .....	107
5.3.4	Effect of altered central inhibition on temporal perception in humans .	108
5.4	Limitations and suggestions for future research .....	109
5.4.1	Open questions and limitations .....	109
5.4.2	Future Experiments.....	110
5.4.2.1	Do humans lose low-SR fibers due to aging? .....	110

5.4.2.2	Are high-SR fibers able to compensate for a loss of low-SR fibers?	110
5.4.2.3	Are speech stimuli more suitable to detect a loss of sensitivity in temporal perception?	111
5.4.2.4	Is age-related change of sensitivity in temporal perception affected by altered central inhibition?	111
5.5	Final conclusions	112
<b>6</b>	<b>Appendix</b>	<b>113</b>
6.1	References of Chapters 1 to 6	114
6.2	List of Figures	137
6.3	List of Tables	138
6.4	Declaration	139
6.5	Danksagung	140

Declaration

## **Abbreviations**

ABR = Auditory Brainstem Response

ANF = Auditory Nerve Fiber

ANOVA = Analysis of Variance

ARHL = Age-Related Hearing Loss

BF = Best Frequency

BM = Basilar Membrane

BMF = Best Modulation Frequency

CAP = Compound Action Potential

CF = Characteristic Frequency

ENV = Envelope

GABA = Gamma-Aminobutyric Acid

GLMM ANOVA = General Linear Mixed-Model Analysis of Variance

H = Harmonic Tone

HHL = Hidden Hearing Loss

HL = Hearing Loss

I = Inharmonic Tone

IC – Inferior Colliculus

IHC = Inner Hair Cell

MTF = Modulation Transfer Function

NIHL = Noise-Induced Hearing Loss

OHC = Outer Hair Cell

SCHR = Schroeder Phase

SCHR+ = upward Schroeder phase

SCHR- = downward Schroeder phase

SE = Standard Error

SNHL = Sensorineural Hearing Loss

SR = Spontaneous Rate

TFS = Temporal Fine Structure

TFS1 = Test of Sensitivity to Temporal Fine Structure

vR = van Rossum

# 1 Introduction

In 2018 the World Health Organization estimated that the world's population with age over 60 years will increase from 12 % to 22 % until 2050 (World Health Organization, 2018). Furthermore, it is estimated that the global number of people suffering from disabling hearing loss (HL) will also double by 2050 from 466 million to 900 million (World Health Organization, 2020). One aspect of (HL) is the loss of sensitivity in temporal perception, which will be the focus of the following thesis,

Temporal perception refers to the perception of variation in acoustic stimuli over time. For example, speech stimuli vary in amplitude and frequency in very short time frames, making temporal perception an important component for communication. Also, other aspects of hearing, such as pitch perception and loudness perception or sound localization are based on temporal perception (Moore, 2014). The auditory system developed a high temporal resolution for perceiving stimulus changes over time (de Boer, 1956; Rosen, 1992; Drullman, 1995). Temporal perception is negatively affected by aging and other types of HL (Füllgrabe and Moore, 2018) and needs in severe cases treatment to maintain the ability of older people to communicate. Untreated HL can also lead to depression (Mener et al., 2013), dementia (Lin et al., 2011) or diabetes (Frisina et al., 2006). Especially in recent years, research on the consequences of HL on temporal perception gained importance. To improve the treatment of impaired temporal perception further knowledge about the physiological origin and the psychoacoustic consequences is required.

As a step to this improvement, the overall aim of this thesis was to investigate the influence of different HL types on behavioral temporal sensitivity in the same animal model, the Mongolian gerbil. Furthermore, conclusions about physiological changes in these animals with different types of HL and their consequences for temporal perception should be drawn from the results. Therefore, temporal perception and its neural correlate at two stages of the auditory pathway were examined in young normal-hearing gerbils. Additionally, the influence of age-related HL and experimentally induced synaptopathy on temporal perception was investigated. The comparison of the results obtained in the gerbil to humans gives further insight into the effects of aging and synaptopathy on temporal perception in humans. The following introduction gives an overview of temporal processing in the auditory system by

focusing on both the peripheral and the central auditory systems. First, the physiological processing of temporal stimuli in the healthy ear is outlined followed by a description of HL- and aging-related consequences for temporal processing. Next, the advantages of the gerbil as an animal model in hearing research and the properties of its auditory system with a focus on temporal perception are summarized. Finally, the thesis aim is outlined with a brief description of the three experimental chapters.

## **1.1 Physiological representation of temporal coding**

To understand how temporal information is encoded knowledge about the structure of the auditory system is required. The mammalian auditory pathway can be separated into a peripheral and central system. Sound first enters the peripheral system through the outer ear, which consists of the pinna and ear canal. When passing the middle ear, the sound wave is transformed into a movement by the three ossicles (malleus, incus, stapes). The resulting impedance adjustment (Helmholtz, 1877; Slama et al., 2010) enables the signal transmission into the cochlea, which is filled with the fluids (Aminoff and Daroff, 2014). In the cochlea, which is part of the inner ear, the mechanical movement is converted into a neural signal by the inner hair cells (IHC) (Pickles et al., 1987; Brugge et al., 2003). The outer hair cells (OHC) provide for high sensitivity and sharp tuning of the basilar membrane (BM) by mechanically and non-linearly amplifying quiet sounds more than loud sounds (Rhode, 1971; Ruggero and Rich, 1991; Ruggero et al., 1992; Brownell, 1983). The auditory nerve fibers (ANF), innervate single IHCs via ribbon synapses, sending the neural signal further to the central auditory system. ANFs are classified in groups based on their spontaneous rates (SR) into low-SR, medium-SR and high-SR fibers (cat: Liberman, 1978; gerbil: Schmiedt, 1989; chinchilla: Ruggero et al., 2000; guinea pig: Furman et al., 2013). In general, high-SRs tend to be associated with low activation thresholds and vice versa (Liberman, 1978; Bharadwaj et al., 2014). The dynamic range differs between fiber types (Sachs and Abbas, 1974; Schalk and Sachs, 1980). Medium- and high-SR fibers start firing at the characteristic frequency (CF) at levels near the hearing threshold and saturate at slightly increased levels in quiet. These two fiber types have only small dynamic discharge ranges and cannot distinguish stimuli well at supra-threshold levels (Liberman, 1978; Costalupes et al., 1984). At frequencies off-CF, medium- and high-SR fibers start firing at higher levels, which allows the processing at supra-threshold levels (Winter and Palmer, 1991). Low-SR fibers, however, respond over a



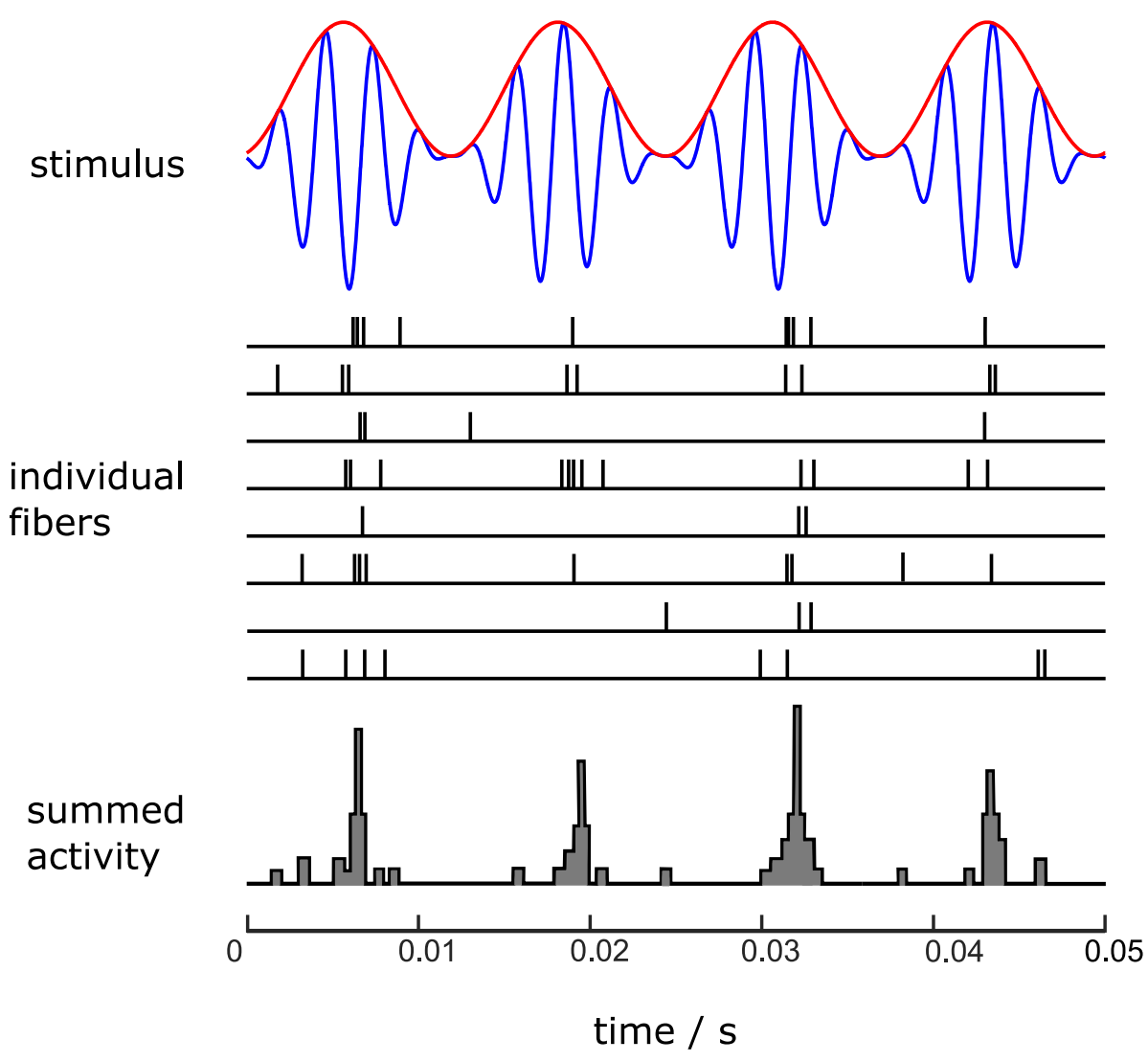
much wider range of sound levels with different discharge rates. Thus, sound levels above hearing thresholds are primarily encoded by low-SR fibers at CF and also by medium- and high-SR fibers off-CF. The low-SR fiber type plays a major role in the present thesis.

The central auditory system consists in ascending order of cochlea nucleus, superior olivary complex, lateral lemniscus, inferior colliculus (IC), thalamus and auditory cortex. In these structures, among other things, the temporal representation of both ears is compared to determine temporal mismatches e.g. for pitch estimation or sound localization (Cariani, 1999).

At the level of the cochlea, temporal features of the acoustic stimulus are encoded mainly by the combination of two measures: Rate-place coding and phase-locking. Due to width and stiffness differences of the BM along the cochlea the peak position in the vibration pattern differs according to the stimulation frequency (von Békésy and Wever, 1960). High frequencies cause a maximum displacement of the BM in the basal region, while the displacement by low frequencies maximizes in the apical region. Exploiting this tonotopy to analyze the frequencies of the signal is referred to as rate-place coding. Thus, each position on the BM behaves like a filter, responding strongly to a limited range of frequencies (Sellick et al., 1982; Khanna and Leonard, 1982). The waveform within each filter can be separated into the temporal fine structure (TFS, rapid oscillation) superimposed by the envelope (ENV, slow amplitude oscillation) (Fig. 1). The filters together form a filter bank, whose center frequency increases approximately in logarithmic step size for frequencies above 500 Hz (Evans et al., 1992; Greenwood, 1990; Glasberg and Moore, 1990). Since ANFs are innervating a single IHC, the stimulus frequencies are reflected by the strongest firing ANFs due to their position along the BM (e.g. Liberman, 2017). The ability to distinguish frequencies by rate-place coding is limited by the filter width since two frequencies exciting the same filter cannot be discriminated. In humans, this leads to a pure-tone frequency difference limen of 0.2 % for the frequency range between 500 Hz and 2000 Hz (Moore, 1973; Sinnott et al., 1992) while the frequency difference limen in gerbils is typically in the range of 10% (Klinge and Klump, 2009).

The second technique to encode temporal features is achieved by synchronized ANF firing and is referred to as phase-locking. Both, phase-locking to TFS and ENV has

been observed (Joris and Yin, 1992). The maximum discharge rate of ANFs is in the range of 300 spikes/s for most mammals including the gerbil (Sachs and Abbas, 1974; Huet et al., 2016). To encode frequencies above 300 Hz, information about the waveform on the BM at the place innervated by a given neuron cannot be carried by single ANFs (Wever and Bray, 1937) but by a group of neurons. Typically, an action potential will not occur on every cycle, but when it is generated, it tends to occur at the same point or phase in the stimulus (Wever and Bray, 1937; Rose et al., 1967). The summed activity of all ANFs reflects the temporal structure of a sound source (Fig. 1). Phase-locking can be quantified by the vector strength of ANF firing and becomes less precise with increasing frequency. For most mammals phase-locking to TFS declines around 1 kHz and becomes close to zero at about 4-5 kHz with some variation across species (cat: Johnsen, 1980; guinea pig: Palmer and Russel, 1986; chinchilla: Temchin and Ruggero, 2010; gerbil: Versteegh et al., 2011), while in the barn owl phase-locking is preserved to some extent up to 10 kHz (Köppl, 1997).



**Figure 1: Schematic representation of phase-locked ANFs in the response of an amplitude modulated pure tone.** Top panel: Amplitude modulated pure tone with TFS in blue and ENV in red. Middle panel: Response of single ANFs to the amplitude modulated pure tone. Bottom panel: Summed ANF activity shows the representation of the ENV periodicity and some representation of TFS. In this way, temporal features of the stimuli are encoded by phase-locking.

In behavioral studies, the limit to benefit from TFS has been used to estimate the upper phase-locking limit in humans. The highest frequency at which TFS information can be used binaurally, e.g. for extraction of interaural time differences, is 1500 Hz (Brughera et al., 2013; Füllgrabe et al., 2017). The human upper phase-locking limit for monaural tasks is still under debate (Verschooten et al., 2013, 2019). An upper limit of 8 kHz was suggested since discrimination in the TFS1 test, a test for TFS sensitivity is still possible up to this limit (Moore and Sek, 2009a, b). In contrast, the lowest suggested upper phase-locking limit of 1.5 kHz for monaural tasks is justified by models showing that no peripheral TFS representation is needed to explain monaural phenomena when correlation in firing rates between cortical neurons tuned to nearby frequencies is assumed. Therefore, it is argued that the limit for binaural TFS perception of 1.5 kHz is also the limit for monaural TFS perception (Verschooten et al., 2019). Recent direct electrophysiological recordings during brain surgery in cochlea nerves of normal-hearing humans confirmed the phase-locking limit of 1.5 kHz (Meijer et al., 2021), which is almost half the limit of 3-4 kHz in gerbils (Versteegh et al., 2011).

Phase-locking to ENV has been observed with amplitude-modulated sounds (Joris and Yin, 1992; Henry et al., 2016). Compared to phase-locking to pure tones, i.e. TFS, the highest modulation frequencies to which phase-locking to ENV occurs is typically lower. Both, TFS and ENV representation are also preserved in the central auditory system with decreasing upper-frequency limit for higher stages of the auditory pathway (Joris, 2003; Devore and Delgutte, 2010). As demonstrated in several mammalian species the upper modulation transfer function (MTF) cutoff frequency to amplitude-modulated stimuli (a measure for ENV representation) in the cochlea nucleus is in the range of 1000 Hz. No effect of degradation can be seen with respect to the ANF (Joris et al., 2004). In the IC, the cutoff frequency decreases for most species down to 100-200 Hz and halves again in the auditory cortex. A measurement of interaural time differences in the cat IC suggested an amplification of ENV coding between the ANF and IC (Joris, 2003). Hence, binaural performance in the central auditory system with increasing frequency may to be dominated by ENV coding rather than TFS coding.

## **1.2 Effects of age-related and noise-induced hearing loss on temporal perception**

HL can have many causes, for example, ototoxic chemicals (Seligmann et al., 1996), oxidative stress (Henderson et al., 2006) or genetic factors. However, the two most prominent contributing factors are noise (noise-induced hearing loss: NIHL) and age (age-related hearing loss: ARHL). The term sensorineural HL (SNHL) is used to refer to pathologies of the sensory cells and neurons within the inner ear (e.g. Kujawa and Liberman, 2015). Both, NIHL and ARHL, can occur due to physiological changes in the peripheral and central auditory system. Several studies have highlighted diminished sensitivity in temporal processing related tasks due to increased audiometric thresholds (e.g. speech perception: Lorenzi et al., 2006; pitch perception: Mathew et al., 2016; gap detection: Lister and Roberts, 2005; binaural perception: Gallun et al., 2014) and ARHL (e.g. speech perception: Plomp, 1989; Peissig and Kollmeier 1997; pitch perception Hopkins and Moore, 2011; gap detection: Snell, 1997; Strouse et al., 1998, binaural perception: Pichora-Fuller and Schneider, 1992; Füllgrabe and Moore, 2018). In the following, the effects of ARHL and NIHL, often in line with elevated audiometric thresholds, and their consequences for temporal perception for humans and animals are summarized. Subsequently, synaptopathy, the dysfunction or loss of synapses in the peripheral system, is described. Finally, the effects of HL originating in the central system and their consequences for temporal perception are outlined.

NIHL can arise gradually from continuous noise exposure, such as background noise, or from sudden exposure to impulse noise (Alberti, 1992). In worldwide collected data 16% of disabling HL was attributed to occupational NIHL (Nelson et al., 2005). Investigation of HL and aging effects are limited in humans to non-hazardous and non-invasive methods. A variety of non-invasive measures have been used to investigate whether noise-induced synaptopathy occurs in humans, but the results were conflicting (review: Bramhall et al. 2019). Relying on electrophysiological or behavioral tasks, evidence was found for noise-induced synaptopathy in humans (e.g. Bharadwaj et al., 2015; Stamper and Johnson, 2015; Liberman et al., 2016; Bramhall et al., 2017) and against (e.g. Spankovich et al., 2017; Grose et al., 2017; Fulbright et al., 2017; Prendergast et al., 2018a and b; Ridley et al., 2018; Guest et al., 2018). Therefore, also animals have to be used relying on invasive studies. NIHL in animals is normally

achieved by artificial short-term noise exposure. These effects may be different from long-term NIHL in humans. For example, a 2 h noise exposure at 100 dB SPL in mice can lead to substantial synaptopathy (Kujawa and Liberman, 2009), but may not have equal consequences for humans, as it is comparable to a rock concert visit.

NIHL in animals was reported in many studies (e.g. Miller et al., 1963; Liberman and Dodds, 1984; Miller et al., 1997) and is often verified by elevated behaviorally evaluated hearing thresholds or altered auditory brainstem response (ABR) thresholds. Depending on the time period of recovery of elevated audiometric thresholds following noise exposure, NIHL has been described as temporary or permanent. Features of temporary threshold shift are not caused by hair cell death but induced due to swelling of cochlea nerve terminals, leading to a temporary disconnection of synapses (Spendlin, 1971; Liberman and Mulroy, 1982). Although the threshold may fully recover, temporary threshold shifts in mice and potentially in other species can lead to acute loss of afferent nerve terminals and delayed degeneration of synaptic ribbons (Kujawa and Liberman, 2009). Noise-induced permanent threshold shifts can have various causes and could involve the destruction of IHCs, OHCs or ANFs. Furthermore, reduced function of the active amplification mechanism in the cochlea resulting in broader tuning of the filters on the BM is often observed (Glasberg and Moore, 1986; Moore, 2007).

The effects of NIHL on phase-locking and thus on temporal coding are different for narrow-band and broadband sounds. NIHL and OHC damage caused by antibiotic treatment (e.g. kanamycin) have either no effect on phase-locking to TFS of sinusoidal stimuli (cat: Miller et al., 1997; guinea pig: Harrison and Evans, 1979) or rather small effects (chinchilla: Woolf et al., 1981; Henry and Heinz, 2012). The investigation of amplitude-modulated stimuli also showed no change in phase-locking to TFS but enhanced phase-locking to ENV for a wide range of modulation frequencies (Kale and Heinz, 2010, 2012). Thus, a loss of compression or a steepening of rate-level functions is most likely due to OHC damage (Heinz and Young, 2004; Moore and Glasberg, 2004) and not caused by the broadening of BM filters. On the other hand, phase-locking to TFS of a broadband stimulus like speech is consistently reduced in individuals suffering from NIHL (Miller et al., 1997; Henry and Heinz, 2013) or OHC loss (Palmer and Moorjani, 1993). Additionally, humans suffering from HL show a greater decrease in sensitivity for temporal-related tasks like tone-in-noise perception than for pure

tone perception in comparison to normal-hearing listeners (Moore, 2008). This effect was also observed in phase-locking to TFS of NIHL chinchillas and attributed to broader tuning and therefore decreased signal-to-noise ratio within each BM filter (Henry and Heinz, 2012). Overall, moderate NIHL reduces temporal coding of broadband sounds, but much less temporal coding of narrow-band sounds (Moore et al., 2014).

ARHL affects about two-thirds of people over seventy years of age (Lin et al., 2011) and also reduces the ability for temporal perception (e.g. Moore et al., 2012). Perceptual symptoms of ARHL in humans include elevated hearing thresholds, loss of high-frequency hearing, recruitment due to damaged OHCs, tinnitus and compromised speech perception (Robinson and Sutton, 1979).

The major knowledge about physiological changes due to ARHL, similar to NIHL, has been derived from animal studies. The volume of the stria vascularis of old animals can be reduced (e.g. Schuknecht and Gacek, 1993; Schmiedt, 1996; Ohlemiller, 2009; Heeringa and Köppl, 2019), resulting in a decreased endocochlear potential (Schmiedt et al., 1996). Broader frequency tuning curves on the BM (Evans, 1975; Khanna and Leonard, 1982; Sellick *et al.*, 1982; Ruggero, 1992) and reduced sensitivity to weak sounds are the consequences. This malfunctioning weakens the frequency selectivity in individuals suffering from ARHL and thus the ability to benefit from TFS perception. (Robles and Ruggero, 2001).

A major consequence of noise-exposure and aging can be a loss or dysfunction of ANFs or ribbon synapses (e.g. noise-exposure: Kujawa and Liberman, 2009; aging: e.g. Schuknecht and Gacek, 1993; Makary et al., 2011; Sergeyenko et al., 2013; Gleich et al., 2016; Liberman and Kujawa, 2017; Fischer et al., 2020) resulting in diminished temporal coding shown by degraded vector strength of phase-locking when background noise is present (Henry and Heinz, 2012). The loss of IHCs caused by aging is minimal (Tarnowski et al., 1991). Comparing the loss of each fiber type, in particular, the low-SR fibers are vulnerable to noise exposure (Furman et al., 2013). ARHL in gerbils also affects low-SR fibers, but only if located in the high-frequency region of the BM (Schmiedt et al., 1996). Recent research on gerbils has shown that temporal coding in single ANFs is not degraded due to aging (Heeringa et al., 2020). This result suggests that the qualitative properties of the remaining ANFs should not be responsible for

ARHL. The loss of ANFs, however, may reduce the peripheral input into the central auditory system, which could lead to reduced sensitivity in temporal perception.

Age-related synaptopathy was also observed in humans. Similar to observations in animals, post mortem histology in elderly humans showed little IHC loss, but a severe loss of peripheral axons of ANFs (Wu et al., 2019). The results suggested that many ANFs are disconnected from their IHCs in the elderly human ear. Furthermore, the degree of HL in elderly humans is well predicted by the loss of OHCs, IHCs and ANFs since ante mortem audiometric thresholds and post mortem OHC, IHC and ANF loss correlate (Wu et al., 2020). Despite clear histological signs of synaptopathy in elderly humans, the consequences for temporal perception are not clear. A recent human psychophysical study found no evidence that age-related cochlea synaptopathy has substantial effects on measures of auditory temporal processing or speech reception in noise (Carcagno and Plack, 2021). Additionally, electrophysiological measures in elderly humans showed no evidence for cochlea synaptopathy in the low-frequency region, where phase-locking is strongest, but in the high-frequency region (Carcagno and Plack, 2020).

Animal studies have indicated that deficits in the central auditory system can affect the representation of TFS and ENV and may be responsible for NIHL and ARHL. In the cochlea nucleus, where inputs from several primary neurons are combined (Joris et al., 1994), reduced phase-locking was observed in noise-exposed animals (Woolf et al., 1981). A loss of central inhibition due to ARHL or NIHL was also proposed to negatively affect temporal coding (Shamma, 1985; Gleich et al., 2003; Caspari, 2008; Khouri et al., 2011). The concentration of the neurotransmitter GABA plays a leading role in temporal coding since impaired temporal resolution in old gerbils can be improved by increasing the GABA concentration in the brain (Gleich et al., 2003). Whether these findings are transferable to research on humans has not been shown. It may be that central deficits have an at least comparable impact on temporal perception than peripheral deficits in HL patients.

The effects of age and noise exposure on temporal perception cannot be clearly separated in human studies (e.g. Gallun et al., 2014). Aged humans are exposed to noise throughout their lifetime adding up to a cumulative effect. Additionally, lifetime noise exposure in humans is difficult to quantify. In mice, an interaction of both HL



causes has been observed, since ARHL was accelerated by early age NIHL (Kujawa and Liberman, 2006). In a review of studies on humans comparing a large sample of binaural TFS test results, the factor age was associated with a larger impact on TFS perception than elevated audiometric thresholds (Füllgrabe and Moore, 2018). This result suggests, that on average, people are more strongly affected by ARHL than by NIHL with respect to temporal perception.

Clinically, HL is mostly detected by elevated hearing thresholds in quiet as pure tone audiometry is a fast and cheap measurement technique. However, about 10% of adults display normal hearing thresholds but still suffer from HL (Parthasarathy et al., 2020) which is referred to as hidden hearing loss (HHL) (Schaette and McAlpine, 2011). Most HHL patients suffer from compromised suprathreshold perception like speech or tone in noise perception (Hind et al., 2011; Ralli et al., 2019) as a consequence of compromised temporal perception (Lorenzi et al., 2006). HHL is often the first noted sign of HL and until now no appropriate treatment has been invented. Kujawa and Liberman (2009) showed synaptopathy (loss of afferent nerve terminals and delayed degeneration of ANFs) can occur due to noise-exposure with complete recovery of hearing thresholds, but suprathreshold attenuated ABRs. Since this observation resembles HHL, cochlea synaptopathy is suggested to be responsible for HHL in animals.

Studies on HHL in young adult human subjects with normal hearing thresholds and different lifetime noise exposure revealed no change in sensitivity with increasing noise exposure in various behavioral and electrophysiological tests (Prendergast et al., 2017a, b). An objection could be that the investigated measures were not sensitive enough to detect synaptopathy in humans or that the parameter for lifetime noise-exposure based on self-report was not appropriate to quantify noise-exposure. The effect of noise-exposure on HHL in humans remains controversial (Ridley et al., 2018; Kohrman et al., 2020). This missing effect could be caused by the different duration of noise exposure (short-term in mice, long-term in humans). Aging, on the other hand, could cause HHL in humans since elderly listeners with normal audiometric thresholds show diminished sensitivity in temporal perception (e.g. Hopkins and Moore, 2011; Marmel et al., 2013; Plack et al., 2014; King et al., 2014). The connection between aging and cochlea synaptopathy in humans has not been proven with behavioral and electrophysiological measures, since studies show inconsistent results

(e.g. Prendergast et al., 2019). A recent study, however, investigating post mortem cochlea histology of ante mortem normal-hearing subjects revealed severe loss of ANFs and OHCs in elderly humans (Wu et al., 2019). This finding provides strong evidence, HHL in elderly humans is caused by synaptopathy.

### **1.3 The Mongolian gerbil as an animal model for the experimental evaluation of temporal perception**

Studies with an animal model are not restricted to non-hazardous and non-invasive methods like in humans. The gerbil as an animal model combines several features allowing to optimally study the effects of HL on temporal perception. Gerbils can be kept in smaller spaces compared to larger animals such as cats or rabbits. The main reason for the selection of the animal model, however, was the comparability to humans. Unlike many other rodents, such as rats or mice, gerbils and humans have similar hearing thresholds for low frequencies (Ryan, 1976), which is particularly important for speech perception. This low-frequency hearing has led to a large body of literature on gerbils in hearing research. Their phase-locking limit of 3-4 kHz is well-known (Versteegh et al., 2011), as well as many aspects of their ANF coding and SR-distribution (e.g. Bourien et al., 2014; Huet et al., 2016). Furthermore, the gerbil's temporal perception has been investigated in several behavioral studies (e.g. Hamann et al., 2004, Leek et al., 2005; Klinge and Klump, 2010) and the temporal coding in electrophysiological studies (e.g. Schulze and Langner, 1997; Gleich et al., 2007; Laumen et al., 2016). Despite broader BM filters by about a factor of 1.7 compared to humans (Kittel et al., 2002), gerbils perform similar to humans in some tasks based on temporal perception, like mistuning detection (e.g. Klinge and Klump, 2009).

Furthermore, the effects of HL on temporal perception can be studied in the gerbil. Their life expectancy of 3 to 4 years allows studying age effects while having a reasonable time period for training in behavioral experiments (Cheal, 1986). Old gerbils suffer from cochlea synaptopathy (e.g. Tarnowski et al., 1991; Schmiedt et al., 1996; Gleich et al., 2016; Heeringa and Köppl, 2019) and reduced binding to GABA (gamma-Aminobutyric acid) receptors in the central auditory system (Kessler et al., 2020). Gerbils are also susceptible to NIHL (e.g. Bone, 1978; Boettcher, 2002). The biggest advantage of the gerbil as an animal model relating to TFS studies is the possibility to study the effect of low-SR fiber loss on perception. The perfusion of the drug ouabain

at the gerbils round window results in degeneration of low-SR fibers in dependence on the applied dose (Bourien et al., 2014; Batrel et al., 2017). Overall the gerbil is well suited as an animal model for investigating temporal perception in respect to aging and low-SR fiber loss. In addition, the observed findings will be, to some extent, transferable to humans due to the described similarities of the structure of the auditory pathway and the features of ARHL and NIHL.

## **1.4 Thesis outline and aim**

The investigation of the consequences of synaptopathy on temporal perception is limited to noninvasive measures in humans. Plack et al. (2016) suggested relying on animal data in the same species comparing electrophysiological and behavioral measures with direct histological measures. The present thesis provides the behavioral results of this comparison. Each of the three upcoming chapters (2-4) describes and discusses an auditory behavioral study investigating temporal perception in the gerbil with a focus on the effects of aging and experimentally induced synaptopathy. In the following, chapters 2-4 are briefly outlined, describing the goal of each experiment.

Chapter 2 explores the temporal perception and coding of Schroeder-phase (SCHR) complexes in the gerbil. These stimuli are harmonic-tone complexes and have played a key role in investigating temporal processing. This chapter, for the first time, compares behavioral and neural responses within the same species. Therefore, average-rate and temporal analyses of responses to SCHR complexes in the ANF and the IC are compared to behavioral discrimination in young normal-hearing gerbils. The behavioral results to this chapter can be used as a benchmark to study age-related SCHR complex perception in chapter 3.

The study in chapter 3 aims to investigate the effects of age on temporal perception in the gerbil by measuring the sensitivity of old gerbils to discriminate SCHR complexes. The procedure and stimulus parameters are equal to those of the study presented in chapter 2. This offers the opportunity to compare both behavioral data sets of young and old gerbils to quantify the amount of ARHL to SCHR perception in gerbils. ABR measurements with click stimuli were conducted to estimate the degree of peripheral HL in the old gerbils.

In chapter 4 the effect of ouabain treatment and age on TFS perception is investigated using the TFS1 test. Harmonic and inharmonic tone complexes are discriminated by young, old, sham-operated and ouabain-treated gerbils. As described before, the treatment with ouabain results in a degeneration of low-SR fibers. Similar to the experiments in chapter 3, the degree of peripheral HL is verified by ABRs to click stimuli.

Chapter 5 summarizes the results of the three preceding chapters and discusses overarching insights gained from this thesis. It proceeds with the identification of important limitations and provides suggestions for future research concerning temporal perception in the gerbil. In particular, the comparability and significance of the results for research in humans will be addressed.

## 2 Chapter

# **Neural processing and perception of Schroeder-phase harmonic tone complexes in the gerbil: Relating single-unit neurophysiology to behavior**

Submitted in March 2021 to European Journal of Neuroscience

**Henning Oetjen\*, Friederike Steenken\*, Rainer Beutelmann,  
Laurel H. Carney, Christine Köppl, Georg M. Klump**

\* Shared first author

### **Author contributions:**

G.M.K., L.H.C., and C.K. designed research; H.O., F.S., and L.H.C. performed research; L.H.C., H.O., and R.B. contributed analytic tools; H.O., F.S., L.H.C., R.B., and G.M.K. analyzed data; H.O., F.S., R.B., L.H.C, C.K., and G.M.K. wrote the paper.

### **Acknowledgments**

This work was supported by the DFG priority program “PP 1608”, the DFG Cluster of Excellence EXC 1077/1 “Hearing4all” (Project ID 390895286), NIH-R01-001641 (LHC), and the Hanse-Wissenschaftskolleg.

## 2.1 Abstract

Schroeder-phase harmonic tone complexes have been used in physiological and psychophysical studies in several species to gain insight into cochlea function. Each pitch period of the Schroeder stimulus contains a linear frequency sweep; the duty cycle, sweep velocity, and direction are controlled by parameters of the phase spectrum. Here, responses to a range of Schroeder-phase harmonic tone complexes were studied both behaviorally and in neural recordings from the auditory nerve and inferior colliculus of gerbils. Gerbils were able to discriminate Schroeder-phase harmonic tone complexes based on sweep direction, duty cycle, and/or velocity for fundamental frequencies up to 200 Hz. Temporal representation in neural responses based on the van Rossum spike-distance metric, with time constants related to temporal fine structure or envelope, was compared to average discharge rates. Neural responses and behavioral performance were both expressed in terms of the sensitivity index,  $d'$ , to allow direct comparisons. At the level of the auditory nerve, highest  $d'$  values were based on temporal fine structure for sweep-direction discrimination, and on average rate or temporal envelope for sweep-velocity discrimination. At the level of the inferior colliculus,  $d'$  values based on average rate were highest for all conditions. However, correlations between neural  $d'$  values and behavioral sensitivity for sweep direction were strongest for both temporal metrics, for both auditory nerve and inferior colliculus. Correlations with behavior for sweep-velocity discrimination were highest for either rate- or envelope-based auditory nerve neural  $d'$ , and for either temporal inferior colliculus neural  $d'$ , although, neural  $d'$  values that were capped at maximal values could affect neural-behavioral correlations.

## 2.2 Introduction

SCHR complexes are harmonic tone complexes with components having a phase relationship that results in a periodic frequency sweep (Schroeder, 1970). Interesting perceptual differences between SCHR stimuli challenge some of the basic premises of hearing science, such as the role of the magnitude and phase spectra of complex sounds. Consequently, SCHR complexes have played a key role in psychophysical investigations of cochlea function and neural coding, including the phase response of auditory filters (e.g., Kohlrausch and Sander, 1995; Oxenham and Dau, 2001, 2004; Carlyon et al., 2017), the effect of the phase spectrum on masking (e.g., Smith et al., 1986; Carlyon and Datta, 1997), and sensitivity to TFS (e.g., Dooling et al., 2002; Drennan et al., 2008).

SCHR complexes can be manipulated by varying the phase spectra of these stimuli, without changing the magnitude spectra. The fast, linear frequency sweeps of SCHR complexes occur within each fundamental period of the stimulus; the sweeps are either upward (SCHR-) or downward (SCHR+), depending on the sign of the phase spectrum (Fig. 2). SCHR complexes with opposite sweep directions thus have time-reversed TFS. The duration of the frequency sweep with respect to the fundamental period determines the stimulus duty cycle. SCHR+ and SCHR- complexes with equal duty cycles have equal long-term spectra and similar acoustical envelopes. To date, the perception of SCHR complexes has been evaluated primarily in birds (Dooling et al., 2002) and humans (Kohlrausch and Sander, 1995; Oxenham and Dau, 2001, 2004; Drennan et al., 2008; Carlyon et al., 2017), whereas neural studies have focused on small mammals (e.g., Recio, 2001; Cedolin & Delgutte, 2010). The present study in the Mongolian gerbil for the first time relates behavior and single-unit responses to SCHR complexes at different levels of the auditory pathway in the same species in order to investigate the physiological basis of perception.

Cochlea mechanical measurements in the base confirmed that SCHR+ complexes elicit peakier patterns of excitation on the BM, as compared to SCHR- complexes (chinchilla: Recio and Rhode, 2000; guinea pig, Summers et al., 2003), consistent with a high-frequency-based cochlea model for the masking differences of SCHR stimuli (Kohlrausch and Sander, 1995). Peakier SCHR+ responses are more vulnerable to cochlea compression, consistent with the reduced responses and weaker masking of

SCHR+ stimuli, emphasizing distinct TFS and envelope (ENV) representations throughout different stages within the auditory pathway. Responses to SCHR+ and SCHR- complexes have been tested in both auditory-nerve (AN) fibers and single neurons in the ventral cochlea nucleus (chinchilla: Recio, 2001; cat: Cedolin and Delgutte, 2010). Recio (2001) reported higher firing rates for SCHR- than for SCHR+ complexes for neurons with high CF (the frequency to which a neuron is most sensitive), e.g. CFs > 3 - 4 kHz, whereas for lower CFs the difference between responses to SCHR+ and SCHR- was diminished. The available BM and neural results are inconsistent with psychophysical findings: psychophysical differences between SCHR+ and SCHR- maskers are already relatively strong above 250-Hz signal frequency (Oxenham and Dau, 2001). Thus, there is a gap in our understanding of responses of low-frequency neurons to low-frequency SCHR complexes.

In this study, we tested whether temporal information contained at the single-unit level is sufficient to explain behavioral temporal discrimination. Furthermore, we obtained neural responses to SCHR stimuli in the AN to investigate the ability of fibers to phase lock precisely to the waveforms of the stimulus. In the IC, where neurons receive convergent inputs from several brainstem areas and potentially across CFs, we investigated whether such integration leads to improved neural discrimination. Finally, we compared perception of SCHR stimuli between gerbils and humans. Therefore, we presented gerbils with the same SCHR complexes in behavioral and neural testing. Neural responses were obtained from two stages of the auditory pathway, the AN and the IC, a central processing hub in the midbrain. Two basic types of discrimination were tested using SCHR complexes similar to those in previous behavioral experiments (e.g., Dooling et al., 2002). First, we investigated behavioral and neural discrimination of SCHR+ and SCHR- complexes, for which the direction of the frequency sweep was reversed (sweep-direction experiment). Second, we studied discrimination of SCHR complexes with different slopes of the phase spectra, for which the sweep velocity within each period (sweep-velocity experiment) and duty cycle differed. Comparisons of behavioral and neural responses within the same species provide a deeper understanding of the processing of harmonic tone complexes and frequency sweeps.



## 2.3 Materials and methods

### 2.3.1 Subjects

Adult Mongolian gerbils (*Meriones unguiculatus*) (age 2.75-12 months) were used for all experiments in this study. All protocols and procedures were approved by the Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES), Germany, permit AZ 33.19-42502-04-15/1990. All procedures were performed in compliance with the NIH Guide on Methods and Welfare Consideration in Behavioral Research with Animals (National Institute of Mental Health, 2002). Animals were obtained from the in-house breeding facility of the University of Oldenburg, from a stock derived from Charles River Laboratories. Single-unit recordings in the AN were carried out in six gerbils (2 M, 4 F, age 7 - 8 months) and in the central nucleus of the IC in eleven gerbils (9 M, 2 F, age 2.75 - 4.5 months). Five gerbils (2 M, 3 F) were behaviorally trained in the sweep-velocity experiment, with two additional male gerbils, for a total of seven gerbils in the sweep-direction experiment. Animals were housed individually or in pairs in EU-Type IV cages. Over the time of testing, their age ranged from 6 to 12 months. For comparison to the gerbil results, four human subjects (3 M, 1 F, aged 23 - 28 years) participated in a psychophysical study using the same stimuli as in the behavioral experiments. The experiments were conducted with the understanding and written consent of each subject following the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Basic hearing sensitivity of all subjects was evaluated prior to data collection. In gerbils, ABRs to clicks (Beutelmann et al., 2015) or chirps (0.3 - 19 kHz, 4.2 ms duration) were measured under anesthesia with a combination of ketamine (135 mg/kg for AN and IC, 70 mg/kg for behavior experiments; Ketamin 10%) and xylazine (6 mg/kg for AN and IC, 3 mg/kg for behavior experiments; Xylazin 2%) in a 0.9% saline solution. ABR needle electrodes were placed either near the midline (vertex and neck, in animals destined for behavior or IC recordings, respectively) or near the mastoid of one side and midline of neck (in animals destined for AN recordings), and a ground electrode was placed on one leg. Recordings were amplified (by  $10^4$  for IC and behavior; by  $10^3$  for AN experiments) and band-pass filtered (0.3 - 3 kHz) using an ISO-80 pre-amplifier (World Precision Instruments, Sarasota, FL, USA) and sampled using a digital signal processor (Hammerfall DSP Multiface II (RME Audio, Haimhausen, Germany; 48 kHz sampling rate) controlled by custom MATLAB software (Mathworks,

Natick, MA, USA). Typically, stimuli were presented at a range of levels, from below threshold up to 90 dB SPL, in steps of 5 - 10 dB, repeated 300 - 500 times. Thresholds from averaged ABR waveforms were detected visually at the outset of all experiments and were similar to the thresholds of normal-hearing young gerbils determined by Laumen and colleagues (2016). During neurophysiological experiments, ABR thresholds were re-checked periodically and the experiment was terminated if threshold deteriorated by more than 30 dB. For human subjects, absolute pure-tone hearing thresholds were tested and did not exceed 10 dB hearing level (HL) between 125 Hz and 8 kHz.

### 2.3.2 Schroeder-phase-complex stimuli

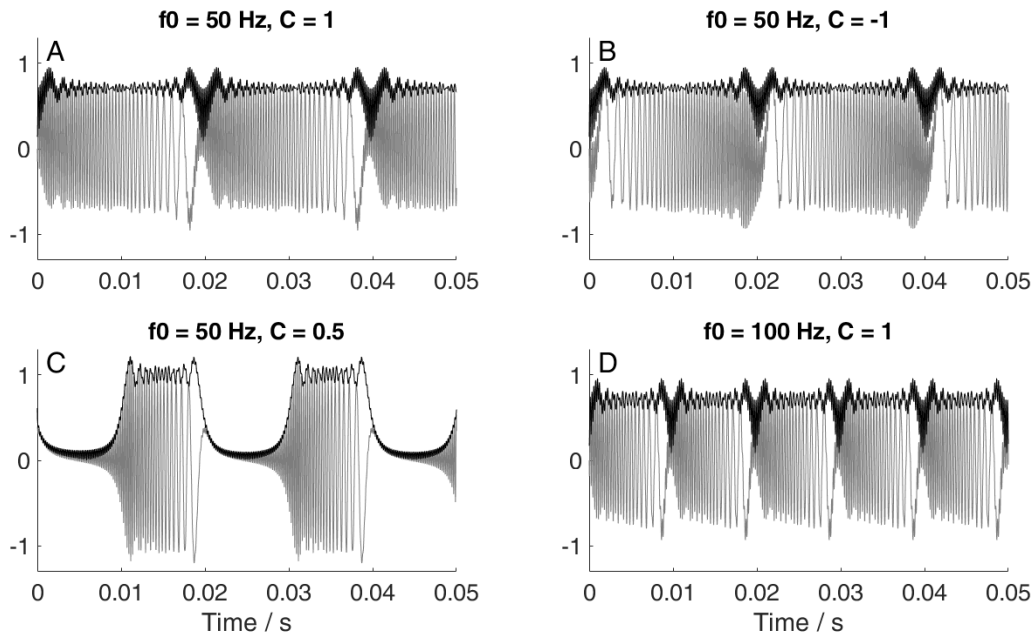
SCHR complexes used in the behavioral and AN experiments were equal-amplitude harmonic-tone complexes with frequencies ranging from  $f_0$  to 5 kHz. The total number of harmonics ranged from 12 to 100 and varied with the value of  $f_0$ , which was 50, 100, 200, or 400 Hz. For IC experiments, stimuli always had 25 components, starting with  $f_0$  (note: despite this unintentional difference between experiments, the SCHR complexes with  $f_0 = 200$  Hz were identical across all experiments). For all SCHR complexes, the phase relationship of harmonics was determined by the following equation:

$$\theta_n = C\pi n(n + 1)/N$$

where  $\theta_n$  was the phase of the  $n^{\text{th}}$  harmonic,  $n$  was the harmonic number, and  $N$  was the total number of harmonics (e.g., see Leek et al., 2005). The absolute value of  $C$  at a given  $f_0$  represents the duty cycle of the frequency sweep (thus the absolute value of  $C$  is inversely related to sweep velocity), and the sign of  $C$  determines the sweep direction (Fig. 2). Values of  $-1 \leq C < 0$  result in an increasing instantaneous frequency sweep, and values of  $0 < C \leq 1$  in a decreasing frequency sweep within each period. Stimuli with  $C = 0$  were pulsatile (no sweep) due to all harmonics being in-phase.  $C$  values of  $\pm 0.25$ ,  $\pm 0.5$ ,  $\pm 0.75$ ,  $\pm 1.0$ , and  $0$  were used. Each stimulus had a duration of 0.4 s, including 25-ms raised-cosine ramps at onset and offset. The overall stimulus level for all SCHR stimuli was 60 dB SPL.

Discrimination of both the sweep-direction and the sweep-velocity experiment were tested behaviorally as well as in responses of single AN and IC units. In all cases, a target stimulus was discriminated from a reference stimulus of the same  $f_0$ . The

sweep-direction experiment investigated discrimination of SCHR complexes that differed only in the sign of  $C$ . All  $C$  values except  $C = 0$  were tested, and both signs served as reference stimuli in different conditions. For SCHR complexes with equal  $f_0$  but sign-reversed  $C$  value, the frequency sweep is time-reversed, but the acoustic envelope and long-term spectra are similar (compare examples in Fig. 2A, B). In the sweep-velocity experiment, the reference was a SCHR complex with a  $C$  value of -1, 0, or 1, and target stimuli had one of the other eight  $C$  values. In the sweep-velocity experiment, discrimination was hypothesized to be based on both acoustic TFS and ENV cues (compare examples in Fig. 2A, C). The TFS and ENV of the acoustic stimulus are altered due to cochlea filtering. For clarity, we refer to the TFS and ENV of the physical stimulus as  $TFS_p$  and  $ENV_p$ , of the cochlea filtered signal as  $TFS_{BM}$  and  $ENV_{BM}$ , and of the AN response to the signal as  $TFS_{AN}$  and  $ENV_{AN}$ . Finally, we refer to these features in the neural response in the IC as  $TFS_{IC}$  and  $ENV_{IC}$ .



**Figure 2: Temporal waveforms of SCHR complexes**, with different values for the fundamental  $f_0$  and the parameter  $C$  which determines the speed of the frequency sweep and, therefore, the duty cycle (absolute value of parameter  $C$ ). Comparison of (A) and (B) shows the effect of the sign of  $C$  on the direction of the frequency sweep. Comparison of (A) and (C) shows the effect of the absolute  $C$  value on the duty cycle. Comparison of (A) and (D) shows the effect of the fundamental on the waveform.

### **2.3.3 Auditory-nerve recordings**

Gerbils were anesthetized intra-peritoneally with a combination of ketamine (135 mg/kg; Ketamin 10%) and xylazine (6 mg/kg; Xylazin 2%) in a 0.9% saline solution. A maintenance dose of one-third of the initial dose was given hourly or when a positive pedal withdrawal reflex was present; the depth of anesthesia was monitored via ECG (obtained with needle electrodes in the right foreleg and the contralateral hindleg), continuously displayed on an oscilloscope. A single dose of non-steroidal antiphlogistic agent (meloxicam, 0.2 mg/kg) was administered subcutaneously. Some animals received pure oxygen (1.5 l/min) via a tube located approximately 1 cm away from their nose. Body temperature (38°C) was controlled via a rectal probe connected to a homeothermic blanket (Harvard Apparatus, Saint-Laurent, Quebec, Canada). Experiments were conducted in a custom-built sound-attenuating booth. The head was held in a bite bar (David Kopf Instruments, Tujunga, CA, USA) and firmly fixed to the setup via a screw glued with dental cement to the exposed skull. The pinna of the right ear was removed to enable placement of a hollow earbar at the bony edge of the ear canal for delivery of the sound. The earbar was then sealed to the ear canal by applying petroleum jelly to obtain a closed sound system. The middle ear was vented by drilling a small hole into the dorsal bulla. At the end of each experiment, the animal was euthanized with an overdose of barbiturate anesthetic (pentobarbital, 48 mg/100 g bodyweight, i.p.).

The AN was accessed via a dorsal approach, by removing the occipital bone and the lateral part of the cerebellum. The brainstem was left intact and was gently pushed medially to access the proximal part of the eighth cranial nerve. A glass electrode (GB120F-10, Science Products GmbH, Hofheim am Taunus, Germany; pulled using a P-2000, Sutter Instruments Co., Novato, CA, USA), filled with 3 M KCl-solution and with a resistance of ~10 - 30 M $\Omega$ , was placed above the AN under visual control and then advanced via a remote-controlled piezo motor (Burleigh 6000 ULN inchworm motor controller and 6005 ULN handset, Burleigh, Inc., Fishers, NY, USA). Recordings were amplified (WPI 767, World Precision Instruments Inc.), filtered for line-frequency noise (50 Hz, Hum Bug, Quest Scientific Instruments Inc., North Vancouver, BC, Canada), and digitized (sampling rate: 48828 Hz; RX6, TDT Inc., Alachua, FL, USA). In parallel, the analog signal was fed to an audio monitor (MS2, TDT Inc.) and displayed on an oscilloscope (SDS 1102CNL, SIGLENT Technologies, Hamburg, Germany).

All sound stimuli were generated by custom-written MATLAB scripts that controlled a digital signal processor (RX6, TDT Inc.) and an attenuator (PA5, TDT Inc.). The signal was then routed through a headphone buffer (HB7, TDT Inc.) and finally played by a miniature earphone (IE 800, Sennheiser, Wedemark, Germany) through the hollow earbar sealed to the ear canal. The sound system was individually calibrated at the start of each experiment via a microphone (ER7-C, Etymotic Research Inc., Elk Grove Village, IL, USA) that was integrated into the earbar, using a microphone amplifier (MA3, TDT Inc.).

Noise search stimuli (1 - 9 kHz at 50 dB SPL, 50 ms duration) were played while the electrode was moved forward into the nerve. When a unit was isolated, its frequency response range was first assessed audio-visually using 50-ms duration tones. The best frequency (BF) of each fiber was then assessed with tones of 50-ms duration, including 5-ms cosine rise/fall times, at a fixed level of 0 - 30 dB above the audio-visually assessed threshold within approximately  $\pm 1$  kHz around the audio-visually determined BF of the fiber. A rate-level function for tones was recorded at BF (10 - 79 dB SPL, 3-dB step size, 10 repetitions, 50-ms duration including 5-ms cosine rise/fall times, 10 repetitions). SCHR complexes, based on a sum of sine functions, were then presented in pairs with opposite signs of  $C$  (SCHR- and SCHR+) at a fixed level of 60 dB SPL, repeated 30 times. A full set of recordings including all  $f_0$  and  $C$  values contained 16 conditions (the SCHR complex with  $C = 0$  was omitted to save recording time). Each SCHR complex pair for the AN recordings, similar to the behavioral stimuli, had a random starting phase in the fundamental period; the phase was kept constant for each physiological dataset.

Final spike detection from the recorded signal was performed off-line, with custom-written MATLAB scripts. Recordings were digitally bandpass filtered (300 - 3000 Hz) and spikes were identified as threshold-crossing events. Importantly, this threshold could be adjusted on a trial-by-trial basis by hand to compensate for variations in spike amplitude. The time of each spike's peak amplitude was saved as the spike time. If more than 10% of the spikes for a given stimulus condition were judged to be below the set threshold criterion due to poor signal-to-noise ratio, that stimulus condition was excluded from further analysis. Furthermore, single-unit isolation was verified by a minimal inter-spike interval of at least 0.6 ms, based on the absolute refractoriness of AN-fibers (Heil et al., 2007). Next, several criteria were checked, designed to exclude

rare units that likely originated in the cochlea nucleus: 1) presence of a pre-potential in the spike waveform, characteristic for anteroventral cochlea nucleus (Keine and Rubsamen, 2015), 2) non-primary-like patterns in the unit's response to pure tones at BF (e.g., Joris et al., 1994; Sinex et al., 2001) and 3) non-monotonic rate-level-functions (e.g., Davis et al., 1996).

#### **2.3.4 Inferior Colliculus recordings**

Gerbils were anesthetized with an initial intra-peritoneal injection of ketamine (Ketamin 10%, 135 mg/kg body weight) and xylazine hydrochloride (Xylazin 2%, 6 mg/kg body weight). A surgical plane of anesthesia was maintained with sub-cutaneous injections (0.03 – 0.05 ml) or slow sub-cutaneous infusion of xylazine hydrochloride and ketamine hydrochloride (0.33 and 7.5 mg/ml, respectively) in 0.9% saline administered with a syringe pump (AL-1000; World Precision Instruments, Sarasota, FL, USA), run at 300–450  $\mu$ l/hr. Sub-cutaneous injections of 0.5 ml of a solution of 0.1 ml atropine (0.5 mg/ml) in 10 ml of 0.9% saline were given at the beginning of the experiment, and again after several hours, to reduce mucous secretions. Lidocaine was used as a topical anesthetic before the initial incision to expose the skull, attach a small headpost to the top of the skull, and place a craniotomy. The animal's body temperature was monitored using a rectal probe and maintained at 38°C by a homeothermic blanket (Harvard Apparatus). Supplemental oxygen was provided through a tube a few cms in front of the animal, as indicated by the blood-oxygen saturation level, which was monitored during experiments (Nonin model 8500AV, Plymouth, MN, USA). Recordings were made with 3 - 4 M $\Omega$  impedance, parylene-insulated tungsten electrodes (Frederick Haer & Co., Inc., Bowdoin, ME, USA) via a small craniotomy located approximately 2 mm lateral to the midline and 0.5 - 1 mm posterior to lambda. The electrode was advanced through the IC using a digitally controlled micropositioner (Burleigh 8200 EXFO inchworm motor controller and 8005 EXFO handset, Burleigh, Inc. Fishers, NY, USA). The neural signals were amplified using an ISO-80 pre-amplifier (World Precision Instruments) and sampled using a Hammerfall digital signal processor (DSP, Multiface II, RME Audio, Haimhausen, Germany; sampling frequency 48 kHz) controlled by custom MATLAB software (Mathworks). The recordings took place in a single-walled sound-attenuating chamber (IAC 401A, Industrial Acoustics, Niederkrüchten, Germany).

Stimuli were generated in MATLAB and presented using the Hammerfall audio interface. Sounds were delivered using earbud headphones (Sennheiser IE 800) through a custom coupler and tube that was sealed into the ear canal using petroleum jelly or earmold material. Stimuli were calibrated in the ear canal using a probe-tube microphone (ER7C, Etymotic Research). All subsequent stimuli were compensated for the magnitude and phase characteristics of the stimulus system.

Each neuron's CF was determined based on responses to 200-ms tones (including 25-ms raised-cos ramps) presented every 600 ms, with levels spanning 10-70 dB SPL and frequencies in a 2-octave range centered on an audio-visual estimate of CF. Neurons were further characterized by a MTF using 500-ms duration, sinusoidally amplitude-modulated tones at CF, with modulation frequencies from 4 - 512 Hz and 5 steps per octave, presented at 60 dB SPL. The dominant peak or valley in the MTF was identified as the best modulation frequency (BMF). SCHR complexes, based on a sum of cosine functions, contained 25 harmonics of  $f_0$  and were presented at 60 dB SPL overall level for 30 repetitions. Each stimulus had the same starting phase within the first fundamental period. All  $f_0$  and  $C$  values were randomly interleaved, and stimuli for each condition were presented in a different random sequence.

At the conclusion of recordings, electrolytic lesions were placed near the recording site (8  $\mu$ A for 25 s, alternating polarity every 5 s). The animals were euthanized with an overdose of barbiturate anesthesia (pentobarbital, 60 – 85 mg/100 g body weight, i.p.) and perfused transcardially with 0.1 M phosphate buffer and 4% paraformaldehyde. Brains were sectioned, stained with cresyl violet, mounted on glass slides, and examined to confirm that lesions were located in the central nucleus of the IC. Based on the relative locations of the lesions, the locations of penetrations, and the depth of each single-unit along its electrode track, we reconstructed the approximate positions of the IC neurons that were studied, in an effort to determine whether there was any obvious clustering or organization of response types within the IC. Other than the well-known tonotopic organization of the IC, no other obvious organization or clustering of MTF or SCHR response types within the IC was observed.

Action potentials were identified off-line using threshold crossings. The threshold for each dataset was determined based on the distribution of spike amplitudes; threshold was set using a well-defined minimum below the peak corresponding to the highest

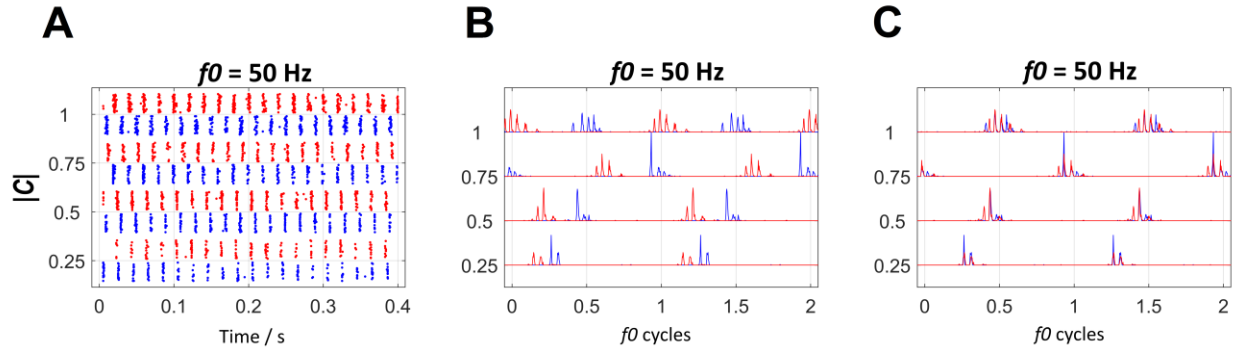
spike amplitudes in the distribution. Only recordings for which there was a well-defined threshold, based on the distribution of spike amplitudes, were accepted as isolated action potentials.

### **2.3.5 Neural recordings - Data Analysis**

The discriminability between two SCHR complexes was estimated based on both average rate and temporal measures applied to the responses of each fiber or neuron. Discrimination based on temporal representation in the responses was quantified using the van Rossum ( $\nu R$ ) distance, which provides a metric for the difference between two spike trains, computed for a given temporal precision (Rossum, 2001). The  $\nu R$  distance increases in proportion to the number of non-coincident spikes between two time-aligned spike trains. The temporal precision of coincidence is determined by the time constant of an exponential tail after each spike, which serves as a coincidence probability weight. For the rate-based analysis, the differences between average rates of spike trains were used as the distance measure.

The SCHR complex is either an upward (SCHR-) or downward (SCHR+) sweep; therefore, a neuron responds when the instantaneous frequency of the SCHR stimulus is near BF (or CF), and thus at different time points of SCHR- and SCHR+ complexes (example in Figs. 2A, B). Direct comparisons of spike trains, without compensation for the different position of BF within the instantaneous-frequency sweep, can lead to large  $\nu R$  distances. To compensate for the time of occurrence of BF within the frequency sweeps, spike trains in response to SCHR- and SCHR+ complexes were analyzed as follows: The first incomplete cycle of the pitch-period was omitted from the analysis. Then the first and last 50 ms were removed from each spike train to avoid onset and offset effects, and a cross correlation of the SCHR- and SCHR+ peristimulus time histograms (bin width 0.1 ms) was calculated. The time lag indicated by the maximum cross correlation was then added to all spike times of the response to one SCHR complex, resulting in both spike trains being aligned with each other (Fig. 3C). Lags larger than one  $f_0$  period were wrapped back into the single period range.  $\nu R$  distances were then computed using the temporally aligned responses.





**Figure 3: Alignment of the spike trains elicited by SCHR+ and SCHR- stimuli prior to the  $\nu R$  analysis.** (A) Raster plots of the response of an AN fiber (BF 1250 Hz, mean SR of  $<1$  spikes per second) stimulated with a 400 ms SCHR complex with a  $f_0$  of 50 Hz. Responses for 30 repetitions each to SCHR- (blue) and SCHR+ (red) complexes are shown, for four different  $C$  values displayed on the ordinate. (B) Period histograms derived from the raster-plots. One cycle is duplicated to clearly visualize the time lag between responses to SCHR+ and SCHR- stimuli which depends on the BF of the fiber. These significant time lags would result in large values of the  $\nu R$  spike distance metric that simply reflect the response delay relative to stimulus onset. (C) To avoid that and probe for differences in the response to the TFS of SCHR- and SCHR+ complexes, the overall relative delay was estimated by cross correlation, and the spike train elicited by one SCHR complex shifted accordingly prior to the  $\nu R$  analysis.

To calculate the  $\nu R$  distances, one group of spike trains was defined as the background or reference condition, the other one as the target condition. In the sweep-direction experiment, the responses to SCHR+ complexes were defined as the reference, those to SCHR- complexes as the target. The time constants for the  $\nu R$  analysis were 1 ms and  $1/f_0$ . Since the resolution of the recorded spike times was  $20 \mu\text{s}$  and the exponential decay for the  $\nu R$  analysis started at the exact spike times, a time constant of 1 ms in the  $\nu R$  analysis is suitable to evaluate the TFS. In the sweep-velocity experiment, AN responses to SCHR complexes with  $C = 0$  were not recorded. Therefore, for AN data, only responses to SCHR- and SCHR+ with  $C = 1$  served as reference, and responses to all other  $C$  values as targets for each  $f_0$ , while for the IC data,  $C = 0$  was also available as a reference. The  $\nu R$  distances between all spike trains of one condition (within and across reference and target repetitions) were saved in a symmetric distance matrix. The diagonal elements (i.e., all self-comparisons) were omitted, because they are zero by definition and distort the result. A subsequent categorization (Escanilla et al., 2015) was carried out to derive a probability of hits and false-alarm rates. A hit was scored if the average  $\nu R$  distance of a target spike train to all other target spike trains was smaller than the average  $\nu R$  distance of that target spike train to all reference spike

trains. In the same manner, a false alarm was scored if the average  $\nu R$  distance of a reference spike train to all target spike trains was smaller than the average  $\nu R$  distance of that reference spike train to all other reference spike trains. The hit and false-alarm rates, that is, the hit and false-alarm counts divided by the total number of target and reference spike trains, respectively, were converted to  $d'$  values by  $d' = z(\text{HR}) - z(\text{FAR})$ , where HR is the hit rate, FAR is the false-alarm rate and  $z()$  is the  $z$ -transform (i.e., the inverse cumulative normal distribution function). To explore the ability of spike times to code the TFS<sub>p</sub> of SCHR complex stimuli,  $d'$  values were estimated for all responses using the  $\nu R$  analysis with a time constant of 1 ms. Furthermore,  $\nu R$  analysis with a time constant of  $1/f_0$  was used to probe for phase locking to the ENV<sub>p</sub>. Thus, for brevity, results obtained with a time constant of 1 ms and  $1/f_0$  are referred to as TFS<sub>AN</sub> or TFS<sub>IC</sub> representation and ENV<sub>AN</sub> and ENV<sub>IC</sub> representation, respectively.

Average discharge rates were calculated from the entire response, including onset. Hits and false alarms were counted in the same way as for the  $\nu R$  distances, using the absolute values of the average-rate difference as a distance criterion.

A  $d'$  higher than 1 represented a discriminable  $\nu R$  or rate difference (Green and Swets, 1966). Values of  $d'$  were limited to 4.3, because they were calculated on the basis of 30 repetitions of each stimulus condition. The effects of the main factors (e.g.,  $C$  value,  $f_0$ ) on  $d'$  were analyzed with a general linear mixed-model analysis of variance (GLMM ANOVA) and, correspondingly, the data are reported using means and standard error (SE). To compare the neural sensitivities with the behavioral outcome, the sensitivities of individual neurons were combined for each condition with the following equation:

$$d'_p = \sqrt{\frac{\sum_i^N (d')^2}{N}}$$

where  $d'_p$  is the pooled sensitivity and  $N$  the number of measured neurons for that condition. Barlow et al. (1971) proposed the lower envelope principle, that behavior is mainly driven by the most sensitive neurons instead of the whole population. To explore this principle, the maximum neural sensitivity  $d'_{\max}$  for each stimulus condition was compared to behavioral sensitivity.

## **2.3.6 Behavioral Experiments**

### **2.3.6.1 Apparatus and stimulus generation: gerbils**

The behavioral experiments with gerbils were carried out in a single-walled sound-attenuating chamber (IAC 401A, Industrial Acoustics) lined with a 15-cm thick layer of sound-absorbing foam (Illbruck Illtec Pyramide 100/50, mounted on Illbruck Illtec PLANO Type 50/0, Cologne, Germany). The reverberation time  $T_{60}$  for broadband white noise was 12 ms, indicating nearly anechoic conditions.

A 30-cm long platform consisting of fine wire mesh was mounted at a height of 90 cm in the middle of the chamber. An elevated wire-mesh pedestal was situated in the center of the platform. The gerbil was trained to sit and wait on the pedestal for target stimuli. Two custom-built light barriers above the pedestal served to monitor the gerbil's position, assuring that the gerbil faced the loudspeaker. The loudspeaker was positioned 30 cm in front of the pedestal at 0° elevation and azimuth relative to the gerbil's head. A custom-built feeder was mounted above the direct sound path and was connected via a flexible tube to a food bowl at the platform. For correct responses, the gerbil received a food reward (a 10-mg custom-made food pellet). The platform, light barriers and feeder produced no relevant sound reflections. The behavioral measurements were carried out without visible light in the chamber. However, the gerbils could be observed under infrared illumination via a closed-circuit video system.

A PC-based Linux workstation with an RME sound-card (Hammerfall DSP Multiface II; sampling frequency 48 kHz) produced the stimuli. The signal was routed through a manual attenuator (Texio type RA-902A, Kanagawa, Japan) to an amplifier (Rotel type RMB 1506, Tokyo, Japan) driving the loudspeaker (Canton Plus XS, Weilrod, Germany; frequency range: 150 Hz - 21 kHz). The absolute sound level at 1 kHz was calibrated every day with a sound level meter (Brüel and Kjaer type 2238 Mediator, Naerum, Denmark) positioned on the elevated pedestal. Stimulus generation, registration of light barriers' switching, and the feeder were controlled by custom software.

### **2.3.6.2 Behavioral procedure: gerbils**

In order to keep the gerbils motivated, the daily amount of food was restricted to maintain a weight of approximately 90% of their *ad libitum* weight, with full access to water.

The gerbils performed in an operant-conditioning Go/NoGo paradigm with food rewards. To ensure that gerbils were facing the loudspeaker when waiting for a target on the pedestal, gerbils were required to start a trial by interrupting the light barriers in the correct sequence. All harmonics of each SCHR complex were based on sine functions. During the entire test session, SCHR complexes were played every 1.3 s as a continuous background of reference stimuli. The starting time within the first fundamental period was chosen randomly for each presented stimulus to avoid potential onset-related cues. After the gerbil jumped on the pedestal, a random waiting time between 1 and 7 s was selected. Jumping off the pedestal before the waiting time elapsed resulted in a restart of the trial with a new waiting time. After the waiting time elapsed, a target stimulus was played instead of the reference stimulus. The target stimuli could either be different from the reference stimulus (test trial) or equal to the reference stimulus (sham trial). Leaving the pedestal (as indicated by the light barrier) within 1 s after target-stimulus onset was registered as a “hit” and rewarded with a 10-mg food pellet. If no response to a target in a test trial was observed, a “miss” was registered, and the gerbil could only initiate a new trial after leaving and re-entering the pedestal. A response during a sham trial was registered as a “false alarm” and a “correct rejection” was registered if the gerbil remained on the pedestal during a sham trial. In the case of a “correct rejection” an additional motivating trial with a salient target was inserted in order to allow the gerbil to obtain a reward. These additional motivating trials were not included in the analysis. For both sweep-direction and sweep-velocity experiments, the initial training started with the expected easiest discrimination. Next SCHR complexes with higher  $f_0$ s were added, and as last step SCHR complexes with lower  $C$  values were introduced. After all conditions were introduced, training continued until the animal achieved a hit rate of at least 50% for the easiest discriminations and a maximum false-alarm rate of 20% in three consecutive sessions. Training typically required about 150 sessions, including the initial training of the waiting period on the pedestal.

Each test session started with a warmup block, which was not included in the data analysis, consisting of eight salient test trials and three sham trials. Next, eight blocks with a randomized order of test trials were presented. All trials within a given block had the same  $f_0$ . In order to interleave simple and more difficult discriminations, blocks of different  $f_0$ s were distributed across each session together with 25 sham

trials. In the sweep-direction experiment, each  $C$  value occurred twice within each block, whereas it occurred only once in the sweep-velocity experiment. Gerbils were tested in the sweep-velocity experiment first and only started with the training for the sweep-direction experiment after completing the former.

### **2.3.6.3 Behavioral procedure: humans**

Stimuli and presentation strategies were similar to those for the gerbil testing. Therefore, only the differences are described in detail here. Stimuli were produced by an RME soundcard (Hammerfall DSP Multiface II; 48 kHz sampling rate) and the output directly driving the headphones (HDA 200, Sennheiser) of the subject placed in a sound-attenuating chamber (Mini 250, Industrial Acoustics). The listener's responses were indicated via a touchscreen (TF1534MC, Iiyama, Hoofddorp, Netherlands) with visual feedback for correct responses. The stimuli in the sessions were presented in the same block design as for the gerbils.

### **2.3.6.4 Behavioral data analysis**

A session was included in the analysis if the gerbil or human finished all trials, and the false alarm rate did not exceed 20%. Data collection continued by adding sessions until the sample size for each data point was at least 20 trials. Because every condition occurred equally often in each condition, the data collection did not favor certain conditions. The behavioral sensitivity  $d'$  was derived based on signal-detection theory as for the analysis of AN and IC data, using hit and false-alarm rates to allow direct comparison of behavioral and neural sensitivity. To avoid infinite  $d'$  values, hit and false alarm rates of 1 or 0 were corrected by  $1/2N$ , with  $N$  being the number of test or sham trials, respectively. The maximum possible  $d'$  values were 4.1 in the sweep-direction experiment and 4.3 in the sweep-velocity experiment, due to different ratios of unique test trials to sham trials in each experiment.

### **2.3.6.5 Neural and behavioral statistical analysis**

All statistical analyses were carried out using SPSS 26 (IBM Statistics, Armonk, NY, USA). The threshold p-value for significance was 0.05. Correlation coefficients report Pearson  $r$  values. All behavioral and neuronal data were analyzed using GLMM ANOVAs (SPSS procedure MIXED) with the sensitivity  $d'$  as the dependent variable and the  $C$  value and  $f0$  as fixed factors. Subsequent additional GLMM ANOVAs included more factors, e.g., a classification of the BF for AN and IC fibers into sets with BF

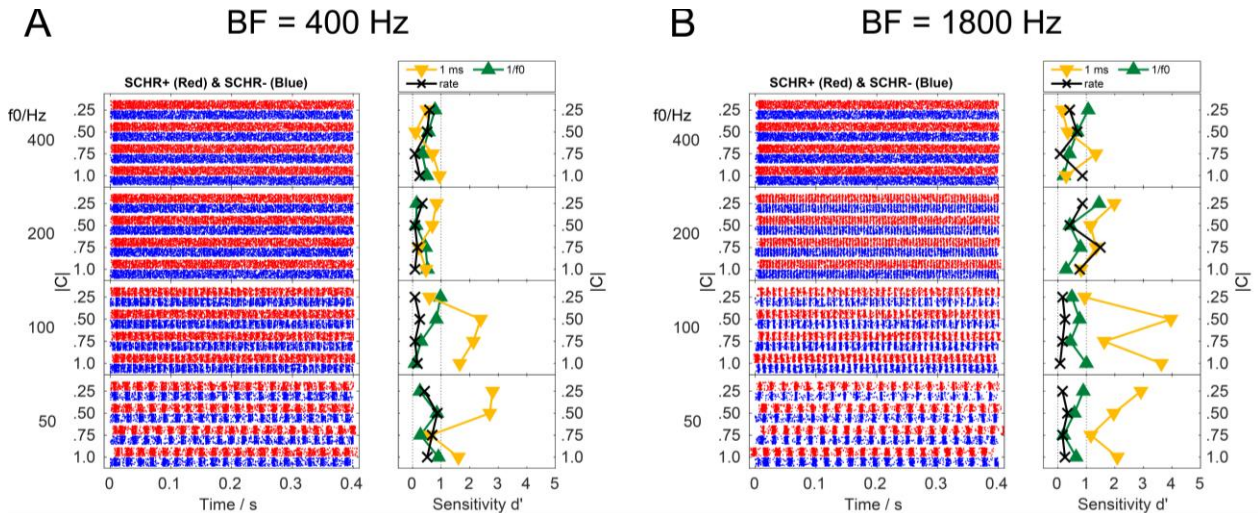
above or below 1.5 kHz. For the IC, the MTF was classified into either bandpass, band-reject, or hybrid MTF type, and the BMF was classified into BMF classes of below 75Hz, between 75 and 150 Hz, or above 150 Hz. Only main effects from the GLMM ANOVA results are reported.

## 2.4 Results

To investigate the representation of SCHR complexes at different stages of the auditory pathway, the ability to discriminate these stimuli based on responses of AN fibers and IC neurons was compared with behavioral discrimination in the gerbil. We distinguished between discrimination experiments providing only acoustic TFS<sub>p</sub> cues (sweep-direction experiment) and those providing both acoustic ENV<sub>p</sub> and acoustic TFS<sub>p</sub> cues (sweep-velocity experiment). In the sweep-direction experiment, discrimination was between SCHR- vs. SCHR+ complexes with  $C$  values only differing in sign, whereas in the sweep-velocity experiment, SCHR complexes with different  $C$  values, could result in different duty cycles, were discriminated. In both experiments neural responses were evaluated by two different  $\nu R$  analyses, using a time constant of 1 ms (referred to as TFS<sub>AN</sub> or TFS<sub>IC</sub> representation) or a time constant of  $1/f_0$  (referred to as ENV<sub>AN</sub> or ENV<sub>IC</sub> representation), as well as an analysis based on mean rate.

### 2.4.1 Neural responses to sweeps that differed in direction

Representative examples of responses from two AN fibers and four IC neurons are shown in Figs. 4 and 5, respectively. The AN fiber in Fig. 4A had BF = 400 Hz, whereas the example in Fig. 4B had BF = 1800 Hz. Both fibers phase-locked to  $f_0$ , as evident in the raster plots, for  $f_0$  values up to 200 Hz. The phase-locking behavior was seen in all fibers tested. The sensitivity index,  $d'$ , for discriminating between pairs of stimuli that differed in sweep direction for the fiber shown in Fig. 4A was higher at low  $f_0$  values (50 and 100 Hz) than at higher  $f_0$  values, when based on TFS<sub>AN</sub> representation (Fig. 4A, yellow). The ENV<sub>AN</sub> (green) and rate (black) representation yielded low  $d'$  values for all SCHR complexes. For the example fiber in Fig. 4B, the  $d'$  based on TFS<sub>AN</sub> representation was comparable to the other example fiber, however, in some cases  $d'$  based on TFS<sub>AN</sub> representation slightly exceeded 1 at  $f_0 = 200$  Hz and  $f_0 = 400$  Hz.  $d'$  values based on ENV<sub>AN</sub> and rate representation exceeded 1 for selected conditions, but were mostly below 1 again. The relative values of  $d'$  obtained with the TFS<sub>AN</sub>-, ENV<sub>AN</sub>- and rate-analyses varied between AN fibers.



**Figure 4: Example responses of two different auditory-nerve fibers (A and B) to positive (SCHR+, red) and negative (SCHR-, blue) SCHR complexes with different  $f_0$  (50 Hz, 100 Hz, 200 Hz, 400 Hz) and  $C$  values ( $\pm 0.25$ ,  $\pm 0.5$ ,  $\pm 0.75$ ,  $\pm 1.0$ ).** The BF of each fiber is given at the top of each panel. Raster plots (to the left of each panel) show responses during the full 400 ms stimulus duration for 30 repetitions. To the right, sensitivity  $d'$  (limited to an absolute value of 4.3) for discriminating SCHR+ and SCHR- complexes are plotted for the same conditions;  $d'$  based on mean discharge rate are shown in black,  $d'$  based on the  $\nu R$  spike distance with a time constant  $\tau = 1$  ms in yellow, and  $d'$  based on the  $\nu R$  spike distance with a time constant of  $1/f_0$  in green. Dot-dashed lines indicate  $d'$  values between 0 and 1. Note that  $d'$  above 1 indicates above-threshold sensitivity. Within each  $f_0$ -panel, responses are arranged in pairs of rows, comparing responses to SCHR+ and SCHR- complexes with the same absolute  $C$  value (i.e., the discrimination made in the behavioral sweep-direction experiment). Sweep velocity increases from the bottom of each panel to the top; that is, the velocity increases with increasing  $f_0$ , and with decreasing  $C$  for each  $f_0$  (see also Fig. 2).

IC neurons had diverse selectivity for sweep direction; the responses in Fig. 5 are for three different MTF types. Fig. 5A shows an example of a neuron with a bandpass MTF that responded vigorously to both SCHR+ and SCHR- complexes across all conditions, with overall rates decreasing at higher  $f_0$  values. Despite the generally vigorous responses, the differences in rates in response to the SCHR- and SCHR+ complexes were reliable, resulting in  $d'$  values for each discrimination that exceeded 1 in nearly all conditions. Note that while the  $d'$  values for the TFS<sub>IC</sub> representation decreased to 0 for  $f_0$  of 200 and 400 Hz, the neural discrimination was still possible based on ENV<sub>IC</sub> representation up to 200 Hz.

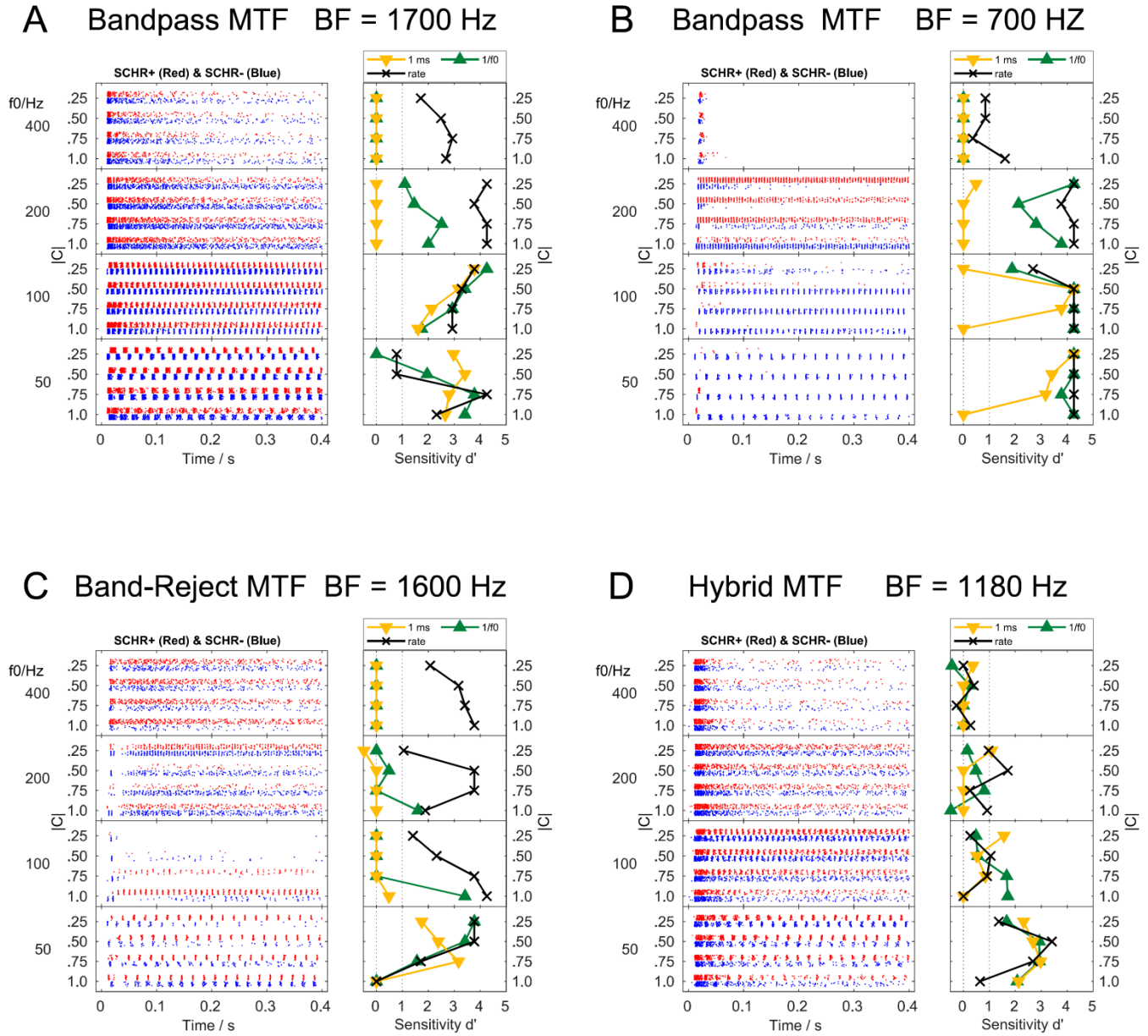
Figure 5B shows a bandpass neuron with striking selectivity for the SCHR complexes that varies with  $f_0$ . It is clear that the responses of this neuron would allow discrimination between the SCHR+ and SCHR- complexes for all conditions with

$f_0 \leq 200$  Hz. Values of  $d'$  were capped at 4.3 (based on the presentation of 30 repetitions), and many conditions reached this cap (e.g., responses to all  $C$  values for  $f_0 = 50$  Hz). Note that the direction of the selectivity for this neuron appears to “reverse” at  $f_0 = 200$  Hz. This apparent reversal may actually reflect a consistent sensitivity of the neuron to fast upward frequency sweeps. A very fast sweep occurs at the end of each fundamental period in SCHR complexes, when the instantaneous frequency “returns” to the starting point of each linear sweep; this fast “return” sweep is in the opposite direction of the main linear frequency sweep.

Fig. 5C illustrates an IC neuron with a band-reject MTF that had reliable rate differences for many stimulus pairs, but with systematic differences in the direction of the difference (recall that stimuli for the entire set of conditions were randomly interleaved during recordings). The band-reject MTF for this neuron had a deep notch at 100 Hz, explaining the generally reduced rates for  $f_0 = 100$  Hz, but note that for  $C = 1$ , the neuron responded strongly to a stimulus even though the  $f_0$  was near the minimum of the MTF.

Finally, Fig. 5D illustrates the responses of a neuron with a hybrid MTF, i.e. responses to amplitude-modulated tones were enhanced with respect to the unmodulated response for modulation frequencies below 100 Hz and were suppressed for modulation frequencies from 300 - 400 Hz. This neuron had robust responses to many of the SCHR conditions, but only had strong selectivity for stimuli with  $f_0$  of 50 Hz, and had weak selectivity ( $d'$  just  $> 1$ ) for some discriminations with  $f_0$  of 100 or 200 Hz. In general, IC neurons could show very strong selectivity for some  $f_0$  and  $C$  combinations, and occasionally for all of the conditions that were tested.





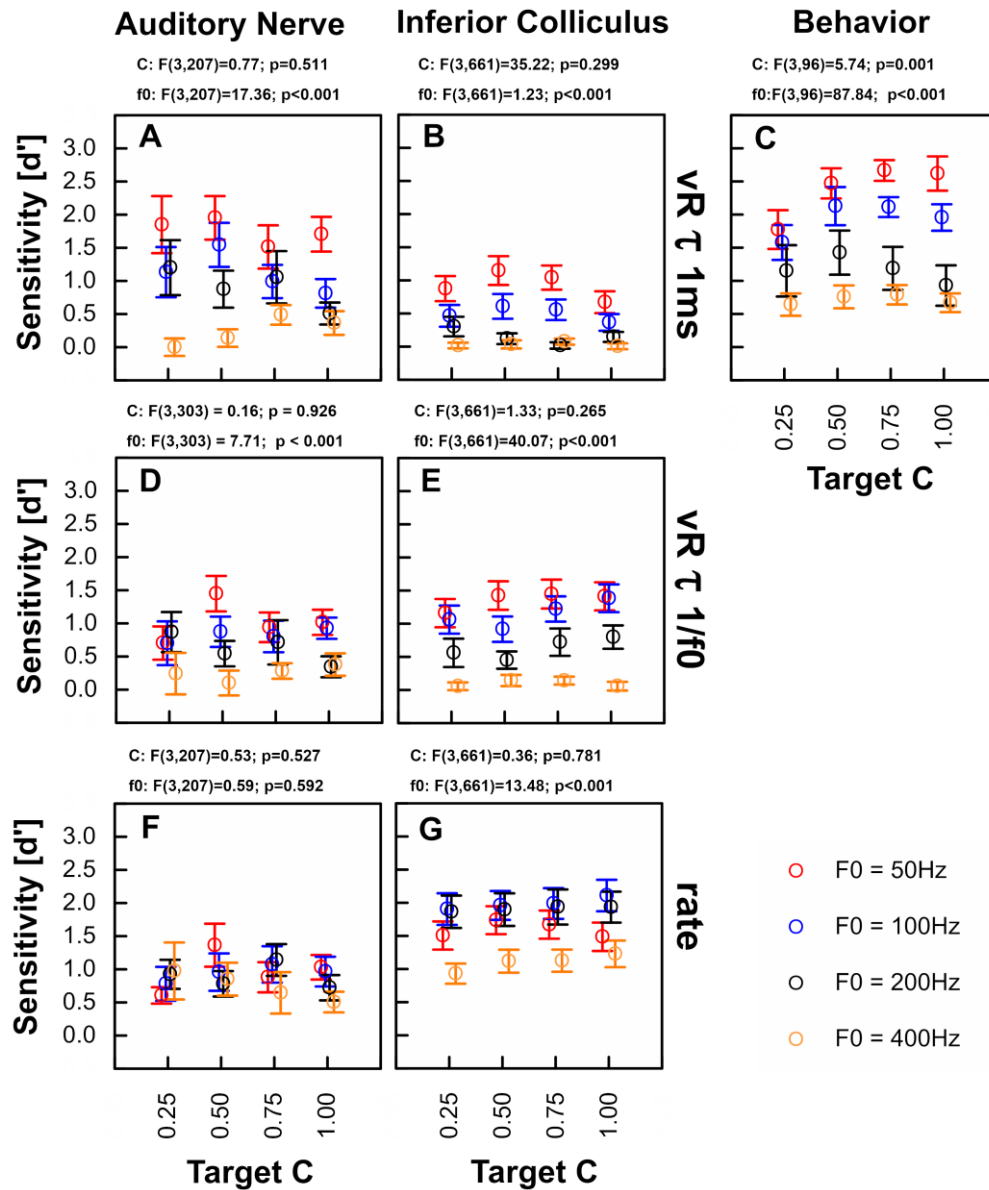
**Figure 5: Example responses of four different neurons (A, B, C, and D, with modulation and frequency tuning indicated at the top of each panel) in the IC. Responses to SCHR+ (red) and SCHR- (blue) stimuli with different  $f_0$  (50 Hz, 100 Hz, 200 Hz, 400 Hz) and  $C$  values ( $\pm 0.25$ ,  $\pm 0.5$ ,  $\pm 0.75$ ,  $\pm 1.0$ ) are shown, in the same layout as Fig. 4.**

#### 2.4.2 Discrimination of sweep direction based on AN responses

Responses of 23 AN fibers from 3 male and 3 female gerbils were analyzed to estimate discrimination performance that could be achieved on the basis of each AN fiber's responses. Note that our sample was deliberately biased towards fibers with BFs that are known to represent stimulus  $TFS_{BM}$  via phase-locking; above 1500 Hz, the strength of phase-locking to tones at BF rapidly decreases in the gerbil (Versteegh et al., 2011). BFs ranged from 400 to 4600 Hz; 14 of 23 fibers had BFs below 3000 Hz, and 9 fibers had BFs below 1500 Hz.

For the  $TFS_n$  representation (Fig. 6A), the mean sensitivity index,  $d'$ , consistently increased with decreasing  $f_0$ . The highest mean  $d'$  values were obtained for an  $f_0$  of 50 Hz (at all  $C$  values). At an  $f_0$  of 400 Hz, mean  $d'$  values fell to chance level, i.e., below 1. The sensitivity index at intermediate  $f_0$ s (100 and 200 Hz) were typically also intermediate. There was no clear dependence of  $d'$  on the  $C$  value. A GLMM ANOVA identified  $f_0$  as the only significant factor that influenced the discrimination sensitivity, whereas neither  $C$  value nor BF class, a factor that was investigated in a subsequent ANOVA, affected  $d'$  values.

For the ENV representation (Fig. 6C), the mean  $d'$  values for nearly all conditions were lower than for the results with a time constant of 1 ms, and almost uniformly below 1. A significant effect of  $f_0$  remained with a tendency towards higher  $d'$  values for lower  $f_0$  values (GLMM ANOVA). Neither  $C$  value nor BF class (in a subsequent GLMM ANOVA) were significant factors. Finally, Fig. 6E displays the mean sensitivity index estimated from AN rate responses to the sweep direction of SCHR complexes. Mean  $d'$  was around 1 for all conditions. A GLMM ANOVA showed no significant effects of either  $f_0$ ,  $C$  value, or BF class, which was tested with a subsequent GLMM ANOVA. In summary, the spiking patterns of AN fibers, analyzed at high temporal resolution ( $TFS$  representation), were best suited to discriminate the sweep direction of SCHR complexes for the lower  $f_0$  values of 50 and 100 Hz (Fig. 6A, C, and E). Discrimination based on  $TFS_{AN}$  representation was best at the lowest  $f_0$  and broke down at an  $f_0$  of 400 Hz. In contrast, both rate-based coding and  $ENV_{AN}$  coding provided poor discrimination of sweep direction at all  $f_0$  and  $C$  values tested. Thus, when the only cue was sweep direction, changes in  $TFS_{AN}$  could explain discrimination, especially for low  $f_0$ s. Changes in rate and in  $ENV_{AN}$  for most comparisons could not explain discrimination.



**Figure 6: Sensitivity  $d'$  (mean,  $\pm$ SE) for discriminating a SCHR complex target stimulus of opposite-sign C value from a reference stimulus, as a function of C value and  $f_0$  (color-coded). Data of this experiment, discriminating sweep-direction, are shown separately for AN fibers (average  $N = 14$ , range 9 to 19; panels A,C, and E), neurons in the IC core (average  $N = 42$ , range 38 to 45; panels B,D, and F) and for the behavioral response of the gerbil ( $N = 7$ ; panel G) and human ( $N=4$ ; panel H). Sensitivity of AN and IC responses is further subdivided according to the response parameter analyzed: (A) and (B) show TFS-related  $d'$  ( $\nu R$  distance with a time constant  $\tau$  of 1 ms), (C) and (D) show ENV-related  $d'$  ( $\nu R$  distance with a time constant  $\tau$  of  $1/f_0$ ). (E) and (F) show  $d'$  based on mean discharge rate. On top of each panel, p-values for main effects of target C value and  $f_0$  determined from a GLMM ANOVA are given.**

### 2.4.3 Discrimination of sweep direction based on IC responses

The results reported here are based on 45 well-isolated single neurons that had BFs below 5.5 kHz. Rate-based and both  $\nu R$ -based analyses were limited to conditions for which the mean response rate was at least 1 sp/sec for at least one of the two stimuli being discriminated. When this rate criterion was not met (in 4 - 8 % of the comparisons), the percent correct was set to 50% (i.e.,  $d' = 0$ ) for that cell in that condition. The mean  $d'$  values for discriminating the sweep direction of SCHR complexes in the IC, based on  $TFS_{IC}$  representation-,  $ENV_{IC}$  representation, or average rate, are shown in Figs. 6B, D, and F, respectively. For  $TFS_{IC}$  representation, only two mean sensitivity values were slightly above the threshold of  $d' = 1$ ; thus, for most conditions, sweep direction could not be discriminated based on  $TFS_{IC}$  representation by IC neurons. Note, however, that some individual neurons were highly sensitive to sweep direction (e.g., Fig. 5A). When the  $\nu R$ -analysis probed for  $ENV_{IC}$  representation,  $d'$  reached slightly higher values than for the  $TFS_{IC}$  representation, especially for  $f_0 = 50, 100, \text{ and } 200 \text{ Hz}$ , with absolute  $C$  values of 0.75 and 1. In contrast to both  $\nu R$ -based analyses, the rate-based sensitivity values were generally above the threshold of  $d' = 1$ , with the maximum mean  $d' = 2.11$  for  $f_0 = 100 \text{ Hz}$  and  $C = 1$ .

The GLMM ANOVA showed that the value of  $C$  had no impact on discrimination by IC neurons, regardless of the temporal integration window. The impact of  $f_0$ , however, was highly significant. Subsequent GLMM ANOVAs, carried out separately for the two  $\nu R$ -based and the rate-based discrimination, included the neurons' BF class, BMF class, and MTF type as factors. The effects (or lack of same) described above for  $f_0$  and  $C$  value were also found in these ANOVAs. For all discrimination strategies, the BF class had no significant impact on sensitivity, suggesting that  $d'$  was not BF-dependent in the IC. For both  $\nu R$ -based analyses, the BMF class [ $\tau = 1 \text{ ms}$ :  $F(2;476) = 9.311, p < 0.001$ ;  $\tau = 1/f_0$ :  $F(2;476) = 8.565, p < 0.001$ ], but not the MTF type, had a significant influence on sensitivity. The pairwise post-hoc comparisons revealed a maximum sensitivity for the BMF class 75-150 Hz. For rate-based analysis, both BMF and MTF influenced the sensitivity [ $F(2;476) = 3.956, p = 0.020$  and  $F(2;476) = 3.557, p = 0.029$ , respectively]. Sensitivity generally increased with increasing BMF class and a pairwise post-hoc comparison showed that the band-reject MTF type had the highest and the hybrid MTF type the lowest sensitivity.

In summary, the sensitivity of sweep-direction discrimination by rate coding was typically above threshold ( $d' > 1$ ) in the IC, while the average  $\nu R$ -based sensitivity was below threshold. When the only acoustic cue was sweep direction, the  $d'$  values were highest for rate and  $ENV_{IC}$  representations, whereas the  $TFS_{IC}$ -based  $d'$  values were rarely higher than 1. Comparing the two stages of auditory processing investigated here, the mean  $d'$  values of IC neurons were thus highest when based on rate representation, whereas AN fibers achieved the highest mean  $d'$  values based on  $TFS_n$  representation.

#### **2.4.4 Behavioral discrimination of sweep direction by gerbils and humans**

The sensitivity for behaviorally discriminating stimuli that differed in sweep direction, as a function of the absolute  $C$  value, is shown in Fig. 6G. Note that because there was no significant difference in sensitivity for datasets for which SCHR+ or SCHR-complexes served as the reference stimulus, we combined datasets based on the absolute value of  $C$  in this analysis. The gerbils' mean  $d'$  sensitivity ranged from  $d' = 0.53$  for  $f_0 = 400$  Hz and  $|C| = 0.25$  to mean  $d' = 2.92$  for  $f_0 = 50$  Hz and  $|C| = 1$ . The GLMM ANOVA showed significant main effects of  $C$  and  $f_0$ . Thus, behavioral discrimination depended on  $C$ , while neural discrimination did not.

Four human subjects were tested with the same stimuli as the gerbils. Average sensitivity to sweep direction in human listeners ranged from mean  $d' = 0.24$  for  $f_0 = 400$  Hz and  $|C| = 0.25$  to mean  $d' = 4.14$  for  $f_0 = 50$  Hz and  $|C| = 0.50$  (Fig. 6H). A GLMM ANOVA revealed a significant main effect of  $C$  and  $f_0$ . A GLMM ANOVA, with data from both species and the same parameters as for the ANOVA above showed that the sensitivity for discriminating sweep direction in humans was significantly higher than in gerbils [ $F(1;144) = 321.22, p < 0.001$ ].

#### **2.4.5 Discrimination of sweep velocity based on AN responses**

The same set of AN recordings was used to test neural discrimination of SCHR stimuli with different values of  $C$  against a reference of  $C = 1$  or  $-1$ . Variation of  $C$  affects both duty cycle and sweep velocity in the stimulus, in addition to the change in sweep direction for a change in the sign of  $C$ . Three fibers were excluded from this analysis because responses to SCHR complexes with  $C = 1$  were not obtained.

We first describe the results for the reference condition with  $C = 1$  (Figs. 7A, D, and F). The mean  $d'$  values based on  $TFS_{AN}$  representation exceeded 1 for stimuli with

$f0 = 50$  Hz and negative  $C$  values (Fig. 7A). For all other  $f0$ , mean  $d'$  values were below 1 for most stimuli, indicating that TFS<sub>AN</sub> representation in the responses to reference and target stimuli did not support discrimination. A GLMM ANOVA confirmed a significant effect of  $f0$  on  $d'$  based on TFS<sub>AN</sub> sensitivity.

For ENV<sub>AN</sub> representation, the  $d'$  values for target stimuli with  $f0$ s of 50 and 100 Hz and  $C$  values of 0.25 and 0.5 were higher compared to the TFS<sub>AN</sub>-based  $\nu R$  analysis. A GLMM ANOVA showed that  $f0$  and  $C$  value both had a highly significant effect on the discrimination of sweep velocity, based on ENV<sub>AN</sub> representation.

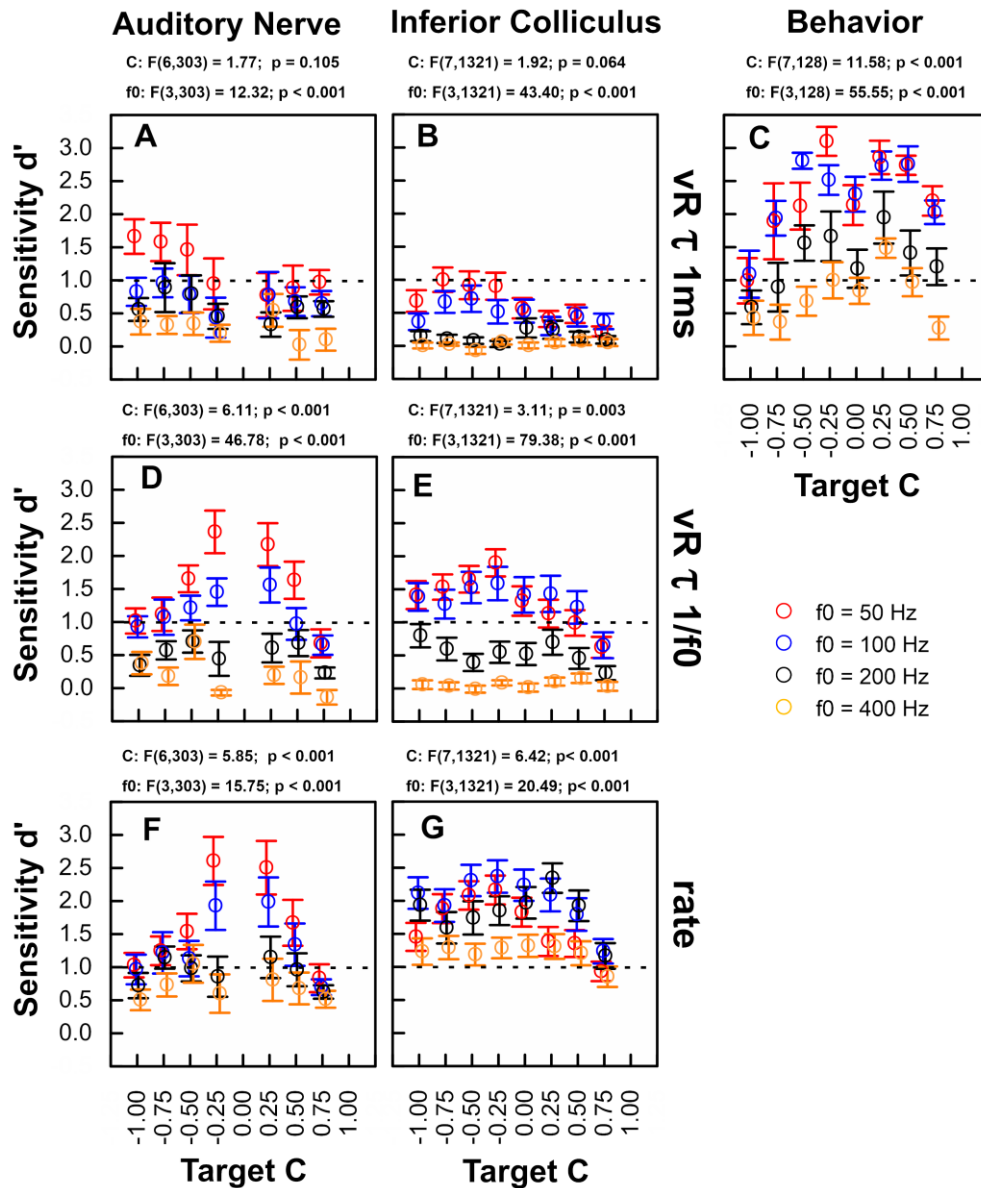
The rate-based analysis, as well as the ENV<sub>AN</sub>-based  $\nu R$  analysis, showed best discrimination for low  $f0$  (i.e.,  $f0 = 50$  Hz and 100 Hz) and  $C = -0.25$  and 0.25 (Fig. 7F). As expected, mean  $d'$  decreased for smaller differences in absolute value of  $C$  between reference and target, i.e., stimuli with more similar duty cycles were harder to discriminate. A GLMM ANOVA showed that rate-based discrimination was affected by  $f0$  and  $C$  value. Additional GLMM ANOVAs for each discrimination type included BF class as a factor and showed no effect of BF class.

Results for the AN-fiber responses analyzed using a reference stimulus with  $C = -1$  (Figs. 8A, D, and F) were similar to those described above for  $C = 1$ . Here, TFS<sub>AN</sub> representation (Fig. 8A) resulted in mean  $d'$  values exceeding 1 for all comparisons of SCHR complex stimuli with  $f0 = 50$  Hz. For  $f0 = 200$  and 400 Hz, most  $d'$  values were below 1, indicating that these discriminations were not possible, based on TFS<sub>AN</sub> representation. Results for  $f0 = 100$  Hz were intermediate. Furthermore, a GLMM ANOVA including BF class revealed that  $f0$  had a significant effect on discrimination ability, based on TFS<sub>AN</sub> representation, whereas  $C$  did not.

ENV<sub>AN</sub> representation (Fig. 8D) showed that  $d'$  values were again higher compared to those based on TFS<sub>AN</sub>-representation, and somewhat lower than those for rate-based discrimination. Both  $f0$  and  $C$  value had an effect on the discrimination of sweep velocity. The additional GLMM ANOVA revealed BF class as a further significant factor for sensitivity [ $F(1,275) = 17.057$ ,  $p < 0.001$ ], such that fibers with BF > 1.5 kHz achieved higher sensitivities, while the highly significant effect of  $f0$  and  $C$  value remained.

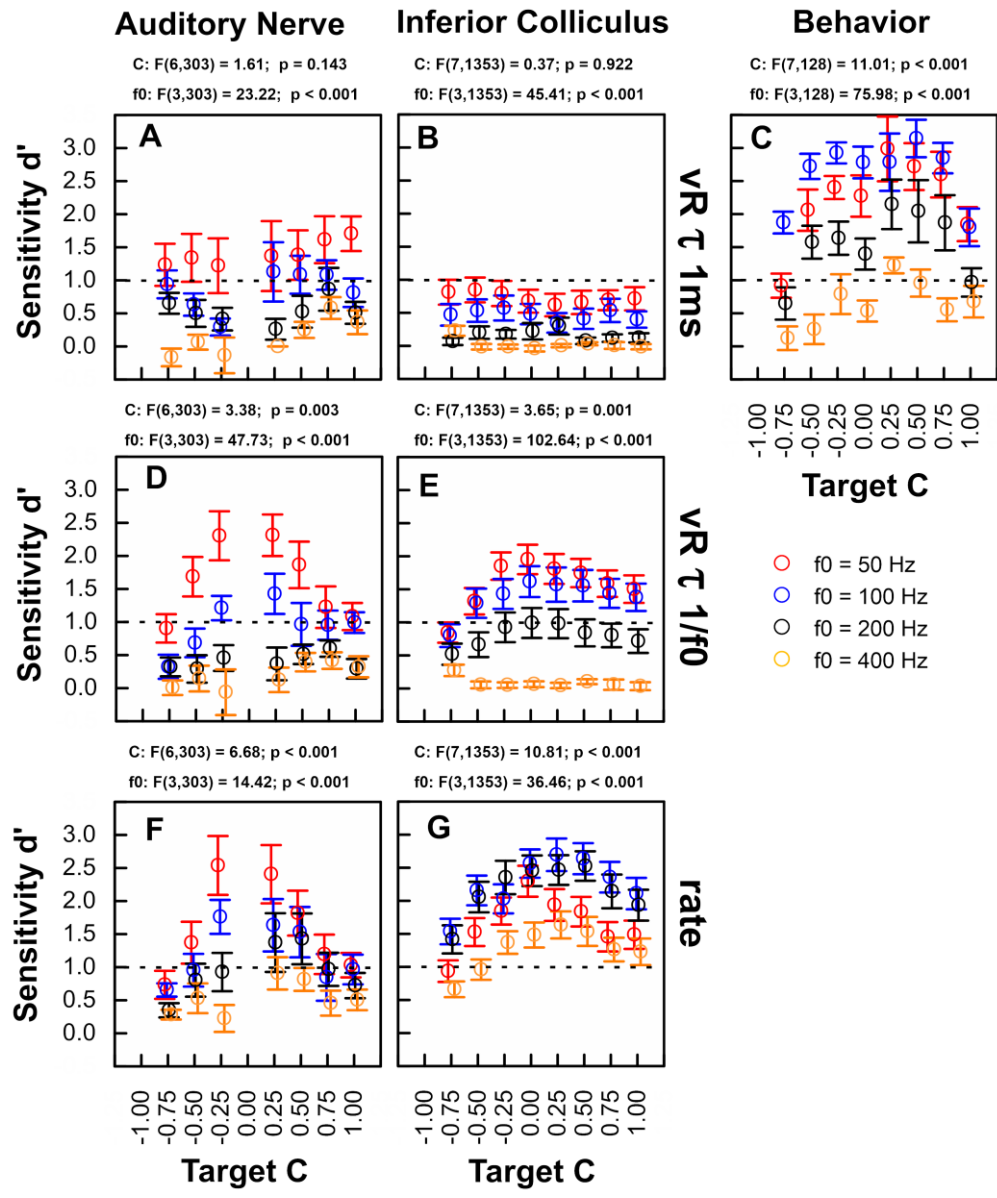
The rate-based analysis (Fig. 8F) resulted in higher mean  $d'$  values than did either of the  $\nu R$ -based analyses, for most comparisons. The highest mean  $d'$  values were again observed for  $f_0 = 50$  Hz. The rate-based discrimination showed a clear dependence on the difference in absolute  $C$  value between reference and target. Mean  $d'$  was maximal when the difference between absolute values of  $C$  was largest, i.e., for target  $C$  values of 0.25 and -0.25, as it was also apparent for  $ENV_{AN}$  representation. Consistent with these qualitative observations, a GLMM ANOVA showed that discrimination based on mean rate was significantly affected by both  $f_0$  and  $C$  value.

In summary, when a velocity cue was present, associated with a change in the duty cycle of the target stimulus, the AN fibers' responses primarily supported discrimination by their  $ENV_{AN}$  representation and mean rate, while the representation of  $TFS_{AN}$  was less salient. Furthermore, presenting opposing sweep directions together with a velocity difference resulted in more discriminable  $TFS_{AN}$  representations than did the velocity cue alone. This difference can be seen in Fig. 7A, for  $f_0 = 50$  Hz and reference  $C$  value = 1, where  $d'$  values were higher for the discrimination with negative  $C$  values, while same-sign discriminations resulted in lower  $d'$ -values. This result is consistent with the results of the sweep-direction experiment and further supports that, in the AN, sweep direction is primarily represented by  $TFS_{AN}$  while the velocity of sweeps and duty cycle are represented by  $ENV_{AN}$  and rate.

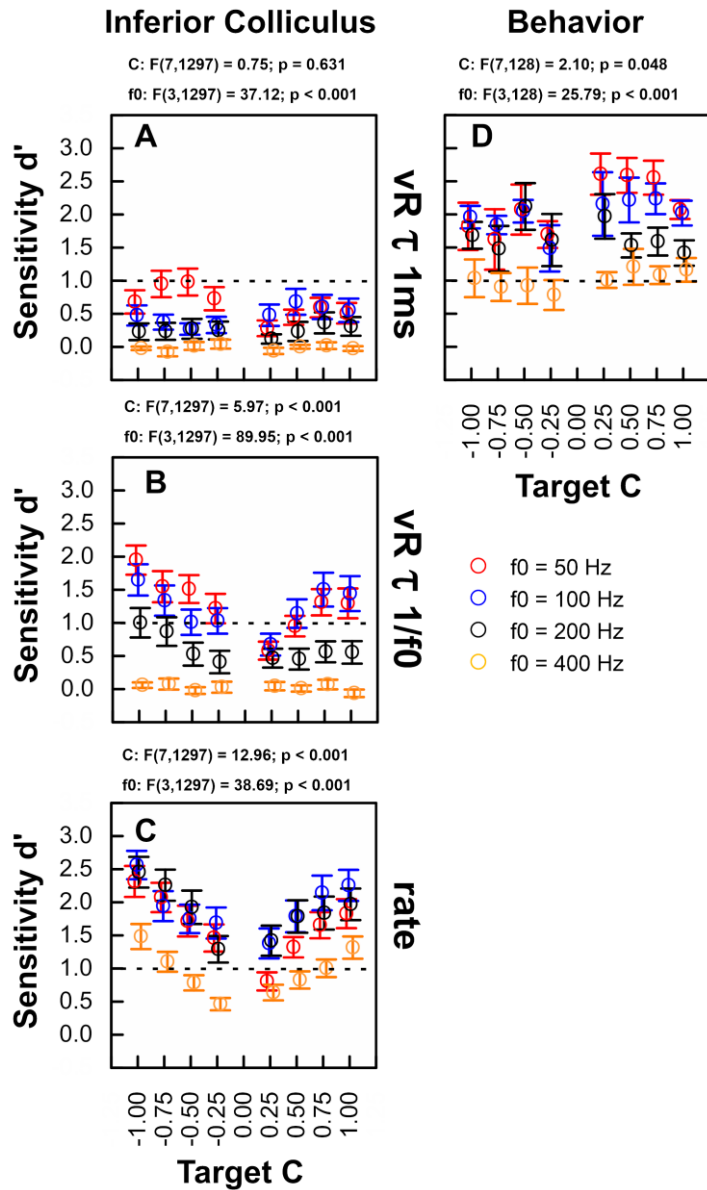


**Figure 7: Sensitivity  $d'$  (mean,  $\pm$ SE) for discriminating a SCHR complex target stimulus from a SCHR complex reference stimulus with  $C = 1$ .** This was one of three reference conditions used to test for discrimination of sweep velocity (continued in Figs. 8 and 9). Each panel plots  $d'$  (mean,  $\pm$ SE) as a function of target  $C$  and  $f_0$  (color-coded). Column panel layout is directly comparable to Fig. 6, with  $C$  values of target indicated on the x-axis. Data shown for AN fibers (average  $N = 12$ , range 5 to 19), neurons in the IC core (average  $N = 43$ , range 38 to 45) and for the behavioral response of the gerbil ( $N = 5$ ). Sensitivity of AN and IC responses are shown separately for the analysis of the  $vR$  distance with a time constant  $\tau$  of 1 ms (TFS-related, A and B), a time constant  $\tau$  of  $1/f_0$  (ENV related, D and E) and spike rates (F and G) together with the gerbils' behavioral sensitivity (C). On top of each subplot, p-values for main effects of target  $C$  value and  $f_0$  determined from a GLMM ANOVA are shown.





**Figure 8: Sensitivity  $d'$  (mean,  $\pm$ SE) for discriminating a SCHR complex target stimulus from a SCHR complex reference stimulus for the reference  $C = -1$ .** Figure layout is the same as Fig.7. Data shown for AN fibers (average  $N = 12$ , range 5 to 19), neurons in the IC core (average  $N = 43$ , range 41 to 45) and for the behavioral response of the gerbil ( $N = 5$ ).



**Figure 9: Sensitivity  $d'$  (mean,  $\pm$ SE) for discriminating a SCHR complex target stimulus from a SCHR complex reference stimulus for the reference  $C = 0$ .** Figure layout is the same as in Fig.7, except no data were obtained for the AN. Data shown for neurons in the IC core (average  $N = 41$ , range 36 to 45) and for the behavioral response of the gerbil ( $N = 5$ ).

#### 2.4.6 Discrimination of sweep velocity based on IC responses

Using the same set of neurons as in the sweep-direction experiment, IC responses were also tested for discrimination of SCHR stimuli with different sweep velocities and duty cycles. Here, the differences in both  $\nu R$ -based analyses and rate-based analyses of IC responses to stimuli with a range of  $C$  values were tested for the ability to support discrimination. Analysis of sensitivity based on both  $\nu R$  analyses and rate are presented separately for reference stimuli with  $C = 1, -1, \text{ and } 0$ .

For the reference  $C = 1$  (Fig. 7), the mean discrimination sensitivity increased with longer integration times. For  $TFS_{IC}$  representation, the best mean sensitivity was  $d' = 1.00$  for  $f_0 = 50$  Hz (Fig. 7B). Compared to that, the  $d'$  values for  $ENV_{IC}$  representation (Fig. 7E), were higher for all  $f_0$  values except for  $f_0 = 400$  Hz. However, despite the increase,  $d'$  values remained mostly below 1 for  $f_0 = 200$ . For  $f_0$  of 50 and 100 Hz, the  $d'$  values were mostly higher for the target  $C$  values with the opposite sign to the reference. Rate-based discrimination resulted in  $d' > 1$  for every  $f_0$ , with a best mean  $d' = 2.37$  for  $f_0 = 100$  Hz (Fig. 7G). Thus, in the IC, discrimination of SCHR complexes with different  $C$  values from the reference  $C = 1$  was generally possible based on  $ENV_{IC}$  and rate representation, but not based on  $TFS_{IC}$  representation. A GLMM ANOVA showed that  $f_0$  affected  $d'$  for every discrimination analysis. The  $C$  value, however, only had a significant impact on  $d'$  based on  $ENV_{IC}$  representation and rate-based discrimination, which also were the only metrics reaching  $d'$  values  $> 1$ . Additional GLMM ANOVAs were carried out for discrimination based on either  $\nu R$ -analyses ( $TFS_{IC}$  or  $ENV_{IC}$  representation) or rate-based analysis, including BF class, BMF class, and MTF type as dependent variables. The results were consistent with the above-mentioned effects for  $f_0$ , but showed no effect for  $C$  value. The BF class only affected rate-based discrimination [ $F(1,957) = 16.49, p < 0.0001$ ], with higher  $d'$  values for neurons with a BF above 1.5 kHz. The modulation-based parameter BMF class had an effect on all metrics ([ $F(2,957) = 21.27, p < 0.001$ ] for  $\tau = 1\text{ms}$ ; [ $F(2,957) = 12.14, p < 0.001$ ] for  $\tau = 1/f_0$  and [ $F(2,957) = 5.46, p = 0.004$ ] for rate). For both  $\nu R$  analyses, the neurons with a BMF of 75-100 Hz reached highest  $d'$  values, whereas a generally increasing sensitivity with increasing BMF class was observed for rate-based discrimination. The MTF class only had an effect on  $TFS_{IC}$  representation [ $F(2,957) = 7.39, p = 0.001$ ], indicating that the band-pass MTF-type neurons could discriminate best between SCHR complexes of different sweep velocity. For the

reference  $C = -1$  a similar pattern as for the reference  $C = 1$  was observed (Fig. 8B, E, and G). The sensitivity also increased with longer integration time and for TFS<sub>IC</sub> representation, no  $d' > 1$  were achieved. Thus, also for  $C = -1$  TFS<sub>IC</sub>-related discrimination of SCHR complexes was not possible (Fig. 8B). ENV<sub>IC</sub> related and rate-based discriminations resulted in sensitivities above threshold (Fig. 8E and G). The trend of lower sensitivity with lower differences between absolute values of reference and target  $C$  can also be seen for these discriminations. The GLMM ANOVAs revealed an effect for  $f_0$  for every discrimination and an effect for  $C$  value for ENV<sub>IC</sub>-related and rate-based discrimination. The additional GLMM ANOVAs showed a BF class dependency only for rate-based discrimination [ $F(1,969) = 5.85, p = 0.016$ ], with neurons with a BF above 1.5 kHz showing higher sensitivity. The BMF class affected the sensitivity for all discriminations ([ $F(2,969) = 21.87, p < 0.001$ ] for  $\tau = 1$  ms; [ $F(2,969) = 9.88, p < 0.001$ ] for  $\tau = 1/f_0$ ; [ $F(2,969) = 13.01, p < 0.001$ ] for rate), reflecting that for ENV<sub>IC</sub>-based  $\nu R$  analysis, neurons with a BMF of 75-150 Hz discriminated best, as it was apparent for discrimination with the reference of 1. The MTF type affected TFS<sub>IC</sub>-based discrimination [ $F(2,969) = 17.05, p < 0.001$ ] and rate-based coding [ $F(2,969) = 3.56, p = 0.029$ ], with band pass MTF types showing the highest sensitivity for TFS<sub>IC</sub> based coding and band-reject MTF type for rate coding.

Finally, for the reference  $C = 0$  sensitivity also increased with longer integration time (Fig. 9A, B, and C). Sensitivities were mostly below 1 for TFS<sub>IC</sub> representation and the sensitivities for ENV<sub>IC</sub> representation and rate decreased with decreasing differences in  $C$  value. The GLMM ANOVAs showed a significant effect of  $f_0$  for all discrimination types and an effect of  $C$  value for ENV<sub>IC</sub> representation and rate. The additional GLMM ANOVAs revealed a main effect for BF class [ $F(1,929) = 8.96, p = 0.003$ ] for rate-based discrimination. The modulation related parameter BMF class influenced the sensitivity for both  $\nu R$ -based discriminations ([ $F(2,929) = 19.48, p < 0.001$ ] for  $\tau = 1$  ms; [ $F(2,929) = 4.83, p = 0.008$ ] for  $\tau = 1/f_0$ ), with BMFs between 75 and 150 Hz obtaining highest  $d'$  values. The MTF class only affected TFS<sub>IC</sub> representation [ $F(2,929) = 7.77, p < 0.001$ ], where the bandpass MTF class reached highest sensitivity values.

In summary, when velocity was added as a cue to SCHR stimuli of opposing sweep direction, spike patterns of IC neurons revealed little TFS<sub>IC</sub> representation, but supported the discrimination primarily by their mean rate, and to a lesser extent by their ENV<sub>IC</sub> representation. Combining both types of change, i.e., cues reflecting sweep

velocity and sweep direction, resulted in higher sensitivities but did not affect the ranking of the suitability of the different neural metrics explored here for discrimination by IC neurons. In contrast, the ranking of neural metrics of AN responses did change when the new cues were introduced in the sweep-velocity experiment.

#### **2.4.7 Discrimination of sweep velocity based on behavioral responses by gerbils**

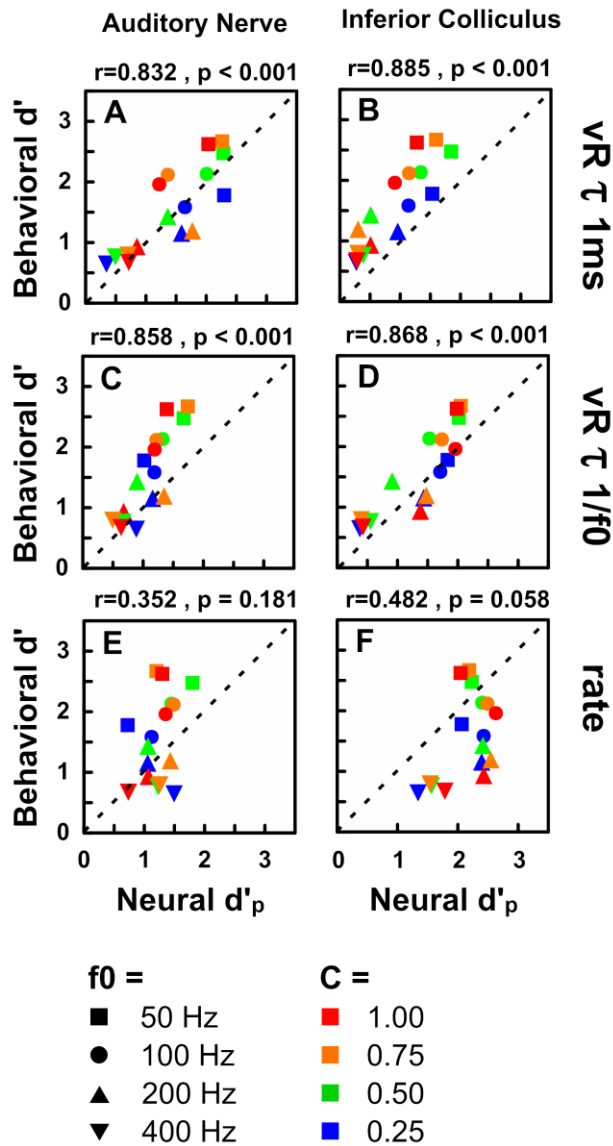
Figures 7C, 8C, and 9D show the behavioral sensitivity,  $d'$  ((mean,  $\pm$ SE) for five gerbils), as a function of the target  $C$  value, for each reference ( $C = 1, -1,$  and  $0$ ). The mean  $d'$  reached 3.1 for  $f0 = 50$  and  $100$  Hz for references  $C = 1$  and  $-1$ , and 2.6 for  $C = 0$ . The values of mean  $d'$  for  $f0 \leq 200$  Hz were above threshold for all three reference conditions. However, for  $f0 = 400$  Hz, most mean  $d'$  values were below 1 for references with  $C = 1$  and  $-1$ , and only marginally above 1 for the reference with  $C = 0$ . For each reference,  $f0$  and  $C$  value were significantly related to  $d'$ . For the references  $C = 1$  and  $-1$ , sensitivities were high for targets with  $C$  values near 0, but decreased for targets with higher absolute  $C$  values, resulting in an inverted U-shape. For the reference  $C = 0$ , however, no particular shape was observed. The general trend was similar to that in the AN and IC: average sensitivity decreased with increasing  $f0$  and decreasing difference of target and reference absolute  $C$  values. The discrimination sensitivity between SCHR+ and SCHR- with  $C = 1$  and  $C = -1$  for  $f0 \leq 100$  Hz were lower in the velocity experiment than in the direction experiment (compare Fig. 6G with 7C and 8C). A possible explanation for this difference is that the training in the velocity experiment caused the gerbils to focus mainly on ENV cues, whereas the training in the direction experiment caused the gerbils to focus mainly on TFS cues.

#### **2.4.8 Correlations between behavioral and neural discrimination**

##### **2.4.8.1 Sweep-direction experiment**

We computed the correlation between the behaviorally obtained  $d'$  values with  $d'_p$  values for the populations of AN and IC responses, for the different combinations of  $C$  and  $f0$ . This analysis was done separately for  $\nu R$ -based analyses with time constants of either 1 ms (TFS<sub>AN</sub> or TFS<sub>IC</sub> representation) or  $1/f0$  (ENV<sub>AN</sub> or ENV<sub>IC</sub> representation) and rate-based analysis (Fig. 10). For AN fibers and IC neurons, both TFS<sub>AN</sub> or TFS<sub>IC</sub> and ENV<sub>AN</sub> or ENV<sub>IC</sub>  $d'_p$  values were highly correlated with the behaviorally obtained

sensitivity (Fig. 10A-D), whereas rate-based  $d'_p$  was not correlated with behavior (Fig. 10E and F). Thus, based on the correlation analysis, differences in rate appear unsuitable to explain the behavioral discrimination of the sweep direction of SCHR complexes. Interestingly, both  $TFS_{AN}$  or  $TFS_{IC}$  and  $ENV_{AN}$  or  $ENV_{IC}$  representations explained the behavior equivalently, despite more prominent acoustical  $TFS_p$  cues for discrimination. In the IC, the  $d'_{max}$  values of more conditions were equal to the maximal possible sensitivity, i.e.,  $d' = 4.3$ , as integration time increased (i.e., from  $TFS_{IC}$  to  $ENV_{IC}$  to rate representations). This saturation of the  $d'_{max}$  values would also have affected the  $d'_p$  values and their correlation with behavior. To test if not only the best neurons, but also an average neuron was sufficient to explain the behavior, the matching to the line of unity was examined. For the pooled AN fibers, the neural sensitivity based on  $TFS_{AN}$  representation and behavioral sensitivities were similar (Fig. 10A), whereas the  $ENV_{AN}$  related  $d'_p$  values were mostly below behavioral  $d'$  values for  $f0$  above 200 Hz (Fig. 10C). For the pooled IC fibers, however, the pattern was reversed, with the  $TFS_{IC}$  representation resulting in lower sensitivities for all  $f0$ s (Fig. 10B) and  $ENV_{IC}$  representation resulting in sensitivities similar to behavior (Fig. 10D). Rate-based coding in both AN and IC did not match the line of unity when compared to behavioral sensitivity. To summarize, correlations between neural  $d'$  values and behavioral sensitivity and matching the line of unity, were strongest for both temporal metrics, for both AN and IC, while rate representation was not correlated with behavior.



**Figure 10: Comparison of mean behavioral discrimination sensitivity,  $d'$ , with neuronal pooled sensitivity,  $d'_p$ , obtained in the sweep-direction experiment.** Correlations for AN fibers ( $N = 16$  conditions; panels A, C, and E) and neurons in the IC core ( $N = 16$  conditions; panels B, D, and F) are shown separately. Furthermore, correlations of the same behavioral data are shown with three different neural measures: temporally based  $d'_p$  of  $vR$  spike distance with a time constant  $\tau$  of 1 ms (TFS-related  $d'$ , panels A and B), a time constant  $\tau$  of  $1/f_0$  (ENV-related  $d'$ , panels C and D) and  $d'_p$  based on mean discharge rate (E and F). Data for different  $f_0$  are distinguished by different symbols, different  $C$  values by color, as indicated in the figure legend. At the top of each panel, Pearson correlation coefficients ( $r$ ) and  $p$  values are given. Dashed lines of unity are included for visual guidance.

#### 2.4.8.2 Sweep-velocity experiment

Analogous to the sweep-direction experiment above, similarities between behavioral  $d'$  and neural  $d'_p$  for the sweep-velocity experiment were assessed using correlations for each reference  $C$  value separately (Fig. 11). All  $p$  values for correlations with  $d'_p$  were lower than 0.05, indicating significant correlations between behavioral and neural  $d'_p$  values, for both AN and IC, regardless of whether the neural discrimination was based on  $TFS_{AN}$  or  $TFS_{IC}$ ,  $ENV_{AN}$  or  $ENV_{IC}$ , or rate representation. Overall, the lowest correlation coefficient was  $r = 0.428$  for  $TFS_{AN}$  representation with behavior, for the reference of  $C = -1$  (Fig. 11B). Correlations between AN neural sensitivity and behavioral sensitivity were higher for rate-based  $d'_p$  than for  $TFS_{AN}$  representing  $d'_p$  values, but similar to  $ENV_{AN}$  representing  $d'_p$  values, irrespective of the sign of the reference. This result differs from the sweep-direction experiment, and suggests that more prominent acoustical envelope cues in the sweep-velocity experiment resulted in rate differences that were large enough to explain the behavioral discrimination.

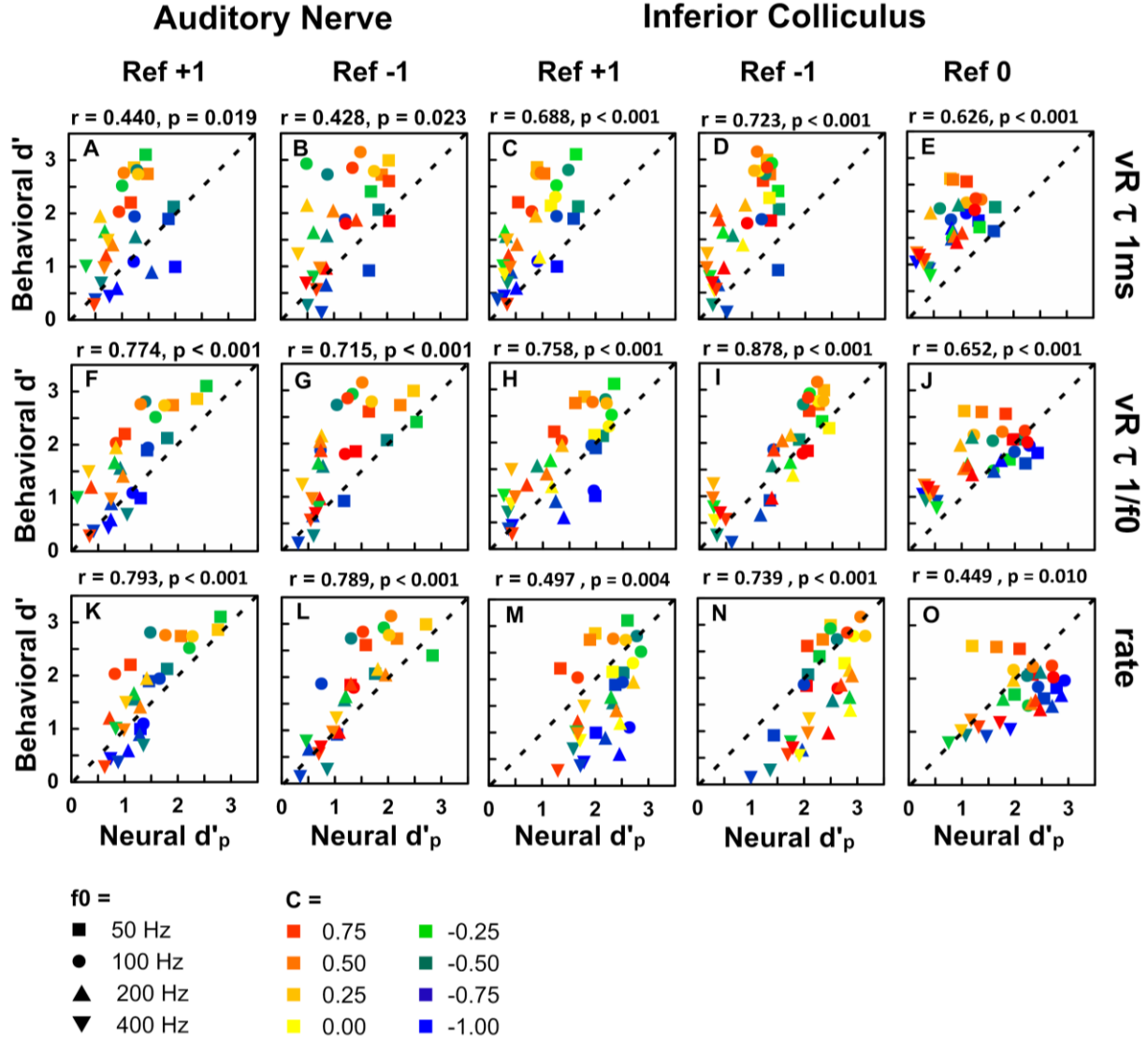
In the IC, pooled  $\nu R$ -based neural sensitivities were correlated better with behavior than were rate-based  $d'_p$  values in two out of three conditions, with the condition  $C = -1$  resulting in similar correlations (compare Figs. 11C, H, M, 11D, I, N, and 11E, J, O). Unlike for the AN, the more prominent acoustic envelope cues in the sweep-velocity experiment did not result in higher correlations between rate-based discrimination in the IC and behavior, compared to both  $\nu R$ -based discriminations. Furthermore, correlations of  $d'_p$  with behavior were similarly high for  $TFS_{IC}$  and  $ENV_{IC}$  representations in the IC (compare Fig 11C and H; 10D and I; 10E and J), unlike in the AN.

The number of conditions with  $d'_{max}$  equal to the maximum sensitivity, i.e.  $d' = 4.3$ , increased in the IC with increasing integration time for all reference conditions. In the AN, however, most conditions did not reach the maximum sensitivity. To determine whether the sensitivity of an average neuron could explain the behavioral results, the proximity of correlations to the line of unity was examined, as in the sweep-direction experiment. In the AN, behavioral  $d'$  only matched neural  $d'_p$  for rate-based discrimination, whereas temporal analyses resulted in mostly lower  $d'_p$  than behavioral  $d'$ . In the IC, however, results strongly depended on the temporal analysis. For all references, the correlation between neural and behavioral sensitivity increased with



integration time, such that the  $ENV_{IC}$ -based sensitivities showed a better match to the line of unity.

To summarize, correlations between neural  $d'$  values and behavioral sensitivity were strongest for the two temporal metrics, for both AN and IC, when only acoustic  $TFS_p$  was present for discrimination (sweep-direction experiment). In contrast, the sweep-velocity experiment included both acoustic  $TFS_p$  and  $ENV_p$  cues and a related change in duty cycle. Here, correlations between behavior and neural  $d'$  values were strongest for either rate- or  $ENV_{AN}$ -based AN analyses, and for both temporal IC analyses, despite the presence of neural  $d'$  values that were capped at maximum values, which likely affected neural-behavioral correlations.



**Figure 11: Comparison of mean behavioral discrimination sensitivity  $d'$  with neuronal pooled sensitivity  $d'_p$  obtained in the sweep-velocity experiment.** Correlations for AN fibers ( $N = 28$  conditions; panels A, B, F, G, K, and L) and IC ( $N = 32$  conditions; panels C, D, E, H, I, J, M, N, and O). Furthermore, correlations of the same behavioral data are shown with three different neural measures: temporally based  $d'_p$  and of  $\nu R$  spike distance with a time constant  $\tau$  of 1 ms (TFS-related  $d'$ , panels A, B, C, D, and E), a time constant  $\tau$  of  $1/f_0$  depending on the fundamental (ENV-related  $d'$ , panels F, G, H, I, and J) and  $d'_p$  based on mean discharge rate (panels K, L, M, N, and O). Panels with the same reference  $C$  value are grouped within on column for AN ( $C = 1$ : panels A, F, and K;  $C = -1$ : panels: B, G, and L) and IC ( $C = 1$ : panels C, H, and M;  $C = -1$ : panels D, I, and N;  $C = 0$ : panels E, J, and O). Data for different  $f_0$  are distinguished by different symbols, different  $C$  values by color, as indicated. At the top of each panel, Pearson correlation coefficients ( $r$ ) and  $p$  values are given. Dashed lines of unity are included for visual guidance.

## 2.5 Discussion

This study explored the potential contributions of temporal coding of both  $ENV_p$  and  $TFS_p$  and of response rate to explaining behavioral discrimination of SCHR complexes. It is useful to distinguish separate issues when comparing  $d'$  for behavioral and neural discrimination: First, which neural response measures provide a possible basis for making the discrimination at the levels of the AN and IC? Second, how well do neural and behavioral discrimination match? Finally, how does perception of SCHR stimuli compare between gerbils, humans, and birds?

### 2.5.1 Mechanisms underlying neural discrimination

Sounds may be represented via rate-place representation and the timing of action potentials in the AN (reviewed by Oxenham, 2018, Huet et al., 2019, but see Carney, 2018). Phase locking encodes both  $TFS_{BM}$  (Rose et al., 1967; Anderson et al., 1971; Johnson, 1980; Joris and Yin, 1992) and  $ENV_{BM}$  modulations (review: Joris et al., 2004). In the gerbil AN, phase locking to  $TFS_{BM}$  extends up to approximately 4 kHz (Versteegh et al., 2011). Phase locking to amplitude modulations typically has a much lower limit,  $<1$  kHz (Dreyer and Delgutte, 2006).

SCHR complexes that differ only in the sign of  $C$  have similar stimulus envelopes and therefore would be expected to be discriminated primarily on the basis of  $TFS_p$  cues. However, cochlea filtering modifies the representation of the stimulus envelope due to dispersive properties of the BM traveling wave and cochlea compression (Smith et al., 1986). These factors predict different ENVs of BM responses such that the down-sweeping SCHR+ complex elicits a more modulated, peakier response than SCHR- (Carlyon and Datta, 1997), as experimentally confirmed for high-frequency BM locations (Recio and Rhode, 2000; Summers et al., 2003). However, as we argue below, these factors do not equally apply to the low-frequency range.

Responses of single AN fibers reflect the BM response of the location they innervate. In the sweep-direction experiment, we thus asked which features of the AN response provide for behavioral discrimination by gerbils. The different  $ENV_{BM}$  shapes and RMS amplitudes of BM responses to the two SCHR polarities (Recio and Rhode, 2000) would be expected to elicit different mean discharge rates in the sweep-direction experiment.

However, a rate difference was not evident (Fig. 6E), plausibly explained by the low-frequency bias of our AN sample. At low frequencies, the dispersive properties of the traveling wave diminish. No direct observations of responses to SCHR complexes are available for apical BM locations; however, diminished differences in response envelopes and amplitudes to SCHR- and SCHR+ complexes would be predicted based on the properties of apical BM and AN responses (Carney et al., 1999; Summers et al., 2003). Consistent with that prediction, a small difference in mean discharge rates evoked by the two SCHR polarities was found in the AN and ventral cochlea nucleus only at CFs higher than 3-4 kHz (Recio, 2001). Our analysis revealed that temporal responses rather than rate responses of AN fibers predicted the gerbils' behavioral discrimination in the sweep-direction experiment (Fig. 10). We conclude that despite known differences in BM response envelopes that arise from cochlea filtering, discrimination of the sweep direction of SCHR complexes with opposite-signed  $C$  values is largely carried by the  $TFS_{AN}$  representation in low-BF AN fibers. Only when additional cues were introduced – velocity and duty cycle, in the sweep-velocity experiment – were the AN rate and  $ENV_{AN}$  representations sufficient to explain gerbils' behavioral performance (Fig. 11). Additional information about the sweep-direction may be derived by patterns of AN responses across fibers tuned to different BFs. For SCHR+, the excitation pattern shifts systematically from base to apex, whereas for SCHR-, AN fibers in the apex are activated prior to fibers in the base. Thus, neurons that receive inputs with a range of BFs can potentially be sensitive to the relative latencies between low and high BF inputs, and thus to sweep direction. Sensitivity of higher-level neurons to latency differences, nevertheless, requires reliable temporal representations in the AN spike times.

Neurons at higher levels of the auditory pathway, as a rule, have weaker phase-locking to acoustic  $TFS_p$ ; however, phase locking in the IC to the input-ENV is stronger than in the AN (Joris et al., 2004). Moreover, temporal features are encoded by derived sensitivities, such as the rate-based MTFs of IC neurons (Schreiner and Langner, 1988; Kim et al., 2020; reviewed by Joris et al., 2004; Walton, 2010). Discrimination of SCHR sweep direction was best transmitted via IC rate. Based on computational-model responses of IC neurons (Nelson and Carney, 2004) to SCHR complexes, it was hypothesized that the sensitivity of IC neurons to the ENV properties of their inputs, as characterized by the periodicity tuning of these neurons (e.g., Joris et al., 2004; Nelson

and Carney, 2007), would be sufficient to explain behavioral sensitivity to sweep velocity and duty cycle. Interestingly, IC responses were in many cases more selective for different SCHR complexes than could be explained based on the neurons' MTFs. For example, Fig. 5 shows responses of IC neurons with rates that were not merely modulated by changes in the parameters of the SCHR stimuli, but that were nearly completely suppressed for some SCHR stimuli, whereas others elicited robust responses. These response differences led to high discrimination sensitivity for both sweep direction and sweep velocity.

### **2.5.2 Relating single-unit responses to behavioral performance**

Parker and Newsome (1998) reviewed two ways of relating neural responses to behavior: the lower-envelope principle (Barlow et al., 1971), in which responses of individual neurons with the highest  $d'$  match the behavioral response, and the pooling principle, in which responses of several neurons are combined to explain behavioral performance. For nearly all combinations of  $f_0$  and  $C$  values, the most sensitive neurons, especially in the IC, were able to discriminate stimuli that were not behaviorally discriminated. For example, behavioral performance was poor for  $f_0 = 400$  Hz, although in many cases neurons had high sensitivity (Figs. 10 and 11). This observation is not unique; it has been reported previously that the sensitivity of individual neurons can be much higher than behavioral sensitivity for the same discrimination (e.g., Klink and Klump, 2004, Johnson et al., 2012, Carney et al., 2014). Such a result, however, contradicts the lower-envelope principle. The pooling principle instead predicts a correlation between a measure of  $d'$  derived from pooling the sensitivity indices of neurons and the perceptual performance. The observed high correlations between neuronal  $d'_p$  and behavioral sensitivity support this prediction. If the neural sensitivity is lower than the behavioral sensitivity, then the responses from a number of neurons must be combined to explain the behavioral sensitivity. Our results indicate that pooling the performance of a small number of neurons (e.g. ~3 AN fibers in the sweep direction experiment; Fig. 10C) would be sufficient to explain behavioral sensitivity. In summary, the high correlations between AN and IC pooled rate- and temporal-based sensitivities and behavior suggest that these neural population responses could be the basis for the behavioral performance, not precluding that additional brain areas (e.g., cortex) are involved.

### **2.5.3 Comparative analysis of perception of SCHR complex stimuli**

Discriminating between sweep direction of SCHR complexes is possible for  $f_0$ s  $\leq$  300 Hz for humans (e.g., Drennan et al., 2008; Lauer et al., 2009, present study), for gerbils (present study), and up to 1000 Hz for birds (Dooling et al., 2002; Lauer et al., 2007). In birds, behavioral performance and compound action potentials (CAP) evoked by SCHR complexes are available in the same species (Dooling et al., 2002). However, differences in CAP amplitudes evoked by SCHR+ and SCHR- stimuli would not explain the animals' discrimination of sweep direction. This is also true for gerbils when comparing CAP measurements reported by Dooling et al. (2002) to the behavioral performance observed in the present study. Thus, CAP peak amplitude is not a valid metric in this context because the CAP is dominated by the highly synchronized onset responses of AN fibers (Özdamar and Dallos, 1978; Bourien et al., 2014). For SCHR-complex stimuli, the temporal synchronization of onset activity across cochlea locations of different CFs dominates the CAP, leading to larger CAP peak amplitudes in response to SCHR- (that resemble upward-sweeping chirps; Dau et al., 2000), compared to SCHR+ (that actually elicit peakier responses at each high-frequency BM location).

Gerbils are often used as an animal model for human peripheral processing due to low-frequency hearing that is comparable to that of humans (Cheal, 1986). Here, we show that both gerbils and humans are able to discriminate behaviorally between sweep direction of SCHR complexes up to  $f_0 = 200$  Hz. However, for low  $f_0$  values, human participants in the current study achieved higher  $d'$  values than gerbils. Gerbils have a shorter cochlea with fewer IHCs (11.2 mm, Müller, 1996; 1400 inner hair cells, Plassmann et al., 1987) and a larger frequency range to represent (0.1 - 60 kHz; Ryan, 1976) as compared to humans (35 mm, Bredberg, 1968; 3500 inner hair cells, Wright et al., 1987; 0.01 - 16 kHz; ISO 389-7, 1996). Thus, the differences in perception between humans and gerbils could be due to differences in the resolution of the representation of SCHR stimuli in the AN.

## **3 Chapter**

# **Aging affects the perception of harmonic tone complexes in the Mongolian gerbil**

### **Author contributions:**

Georg M. Klump designed the study. Henning Oetjen carried out the experiments. Henning Oetjen and Georg M. Klump analyzed and interpreted the data. Henning Oetjen wrote the chapter.

### 3.1 Abstract

Age-related deficits in temporal processing have been observed in humans and other species. Schroeder-phase harmonic tone complexes are suitable for studying temporal processing deficits. They feature periodical linear frequency sweeps with the period depending on the fundamental frequency. A second parameter affects the direction of the sweep and the duty cycle. Here, behavioral discrimination's of Schroeder-phase complexes of different sweep direction for fundamental frequencies of 50, 100, 200, and 400 Hz and duty cycles ranging from 25 to 100 % were studied for three old gerbils (age > 36 months) and compared to the performance of young gerbils (age < 15 months) from a previous study. Auditory brainstem responses to click stimuli in the old gerbils were measured to indicate hearing loss and synaptopathy. Old gerbils were able to discriminate Schroeder-phase complexes differing in sweep direction for fundamental frequencies up to 200 Hz with a decline of performance with increasing fundamental frequency. An age-related decline in sensitivity for old compared to young gerbils was only observed for fundamental frequencies up to 100 Hz. The effect of duty cycle on discrimination sensitivity showed only a trend for old gerbils, possibly due to the small sample size. The auditory brainstem responses of the old gerbils indicated peripheral hearing loss or synaptopathy. A comparison to humans showed a similar age-related decline in sensitivity for discriminating Schroeder-phase complexes. The observed decrease in sensitivity with increasing  $f_0$  also suggests that central processing deficits could have affected perception. An age-related decline of GABAergic inhibition may explain the age-related differences in the processing of Schroeder-phase complexes observed in the gerbil.



## 3.2 Introduction

A major problem in elderly humans is impaired speech in noise reception (e.g. Gordon-Salant and Fitzgibbon, 1993; Stuart and Phillips, 1996; Versfeld and Dreschler, 2002). When testing old participants twice 5 years apart also a decline in speech understanding in quiet and with interfering talkers or reverberation was observed (Divenyi et al., 2005). This decline in speech in noise reception due to aging is generally attributed to deficits in temporal processing (e.g., Hopkins and Moore, 2009; Lorenzi et al, 2006). Age-related deficits in temporal processing have also been found in gap detection (Snell, 1997) and forward masking (Sommers and Gehr, 1998). Schroeder phase (SCHR) complexes, which are harmonic tone complexes with components having a specific phase relation that result in a periodic linear frequency sweep have previously been used for investigating age-related change in temporal perception. Lauer et al. (2009) reported lower performance for discriminating SCHR complexes in old humans compared to young humans indicating that the discrimination of SCHR stimuli suitable paradigm for investigating age-related temporal processing deficits. The present study focuses on the age-related change in temporal perception using the SCHR complexes in the Mongolian gerbil (*Meriones unguiculatus*).

Human studies observed a decline in temporal processing ability with increasing HL (Füllgrabe and Moore, 2017; Gallun et al., 2014) and age (Füllgrabe et al., 2018). It is difficult, however, to separate the effects of age from the effects of HL. Thus, to minimize the influence of HL many studies compared young and old listeners with normal or matched audiograms and still observed a strong effect of age on temporal processing ability (John et al., 2012; Füllgrabe et al., 2015). In recent years, monaural and binaural tests for TFS perception have been developed for investigating temporal sensitivity. The monaural TFS1 test (Moore and Sek, 2009) relies on discrimination of harmonic and inharmonic tone complexes. The binaural TFS-LF test (Hopkins and Moore, 2010) is based on interaural phase shift detection of low-frequency pure tones. Both the TFS1 test and the TFS-LF test (Moore et al., 2012), and a speech in babble noise test (Füllgrabe et al., 2015) confirmed the age-related decline in temporal processing humans. Füllgrabe and Moore (2018) conducted a meta-analysis with the data from 19 studies using the TFS-LF test and confirmed the effect of both age and audiometric threshold on temporal perception. The correlation of temporal processing

sensitivity with age was higher than with audiometric threshold indicating the importance of age per se on temporal perception.

The Mongolian gerbil has become a suitable animal model for aging research having a short lifespan of 3 to 4 years (Cheal, 1986). Especially at low frequencies, where humans benefit from temporal processing for speech perception, the gerbil's low-frequency hearing ability provides for a useful comparison with humans (Ryan, 1976). Age-related temporal processing deficits in the gerbil auditory system have been observed in ABRs elicited by noise bursts separated by silent gaps (Boetcher et al., 1996). Laumen et al. (2016) confirmed the age-related decline in the ABR presenting monaural and binaural click stimuli with different interaural time and level differences. Contrary to the ABR results, Heeringa et al. (2020) recording responses of single ANFs in gerbils of different ages concluded that the auditory-nerve fiber responses cannot account for age-related temporal processing deficits. However, at a higher level of the gerbil auditory pathway, the IC, Khouri et al. (2011) demonstrated an age-related decline in neuronal temporal selectivity. This result and the evidence for a decline in GABAergic inhibition in old gerbils (Gleich et al., 2003; Kessler et al., 2020) suggests that central processing deficits may be important for an age-related decline in temporal perception. Similar to the observations in human subjects (Snell, 1997). Hamann et al. (2004) observed increased gap detection thresholds of old gerbils. Also, forward masking experiments demonstrated a decline in temporal perception both in gerbils and humans (Gleich et al., 2007; Strouse et al., 1998). SCHR complexes have been widely used in animal studies on temporal perception (Dooling et al., 2002; Lauer et al., 2007; Leek et al., 2005) and have revealed age-related temporal perception deficits in humans (Lauer et al., 2009).

The present study investigates SCHR perception by old gerbils and compares the results to those obtained from young gerbils using the same stimulus parameters and procedure (chapter 2). The two parameters which characterize the temporal properties of SCHR complexes are the fundamental frequency  $f_0$ , determining the period of the waveform of the harmonic tone complex, and the  $C$  value, determining the duty cycle within each period and the direction of the frequency sweep (sign of  $C$ ). The comparison of the perception of SCHR complexes by young and old animals allows for a deeper understanding of age-related temporal processing deficits in the Mongolian gerbil.

## **3.3 Materials and methods**

### **3.3.1 Subjects**

Data from three (one male, two female) Mongolian gerbils older than 36 months were obtained. They were food deprived to around 90% ad libitum weight to keep them motivated during the behavioral experiments. They had unlimited access to water. Outside experimental sessions they were housed in a Type IV cage with nesting material. All protocols and procedures were approved by the Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES), Germany, permit AZ 33.19-42502-04-15/1990. All procedures were performed in compliance with the NIH Guide on Methods and Welfare Consideration in Behavioral Research with Animals (National Institute of Mental Health, 2002).

### **3.3.2 Auditory brainstem response measurement**

To evaluate the hearing ability of the old gerbils the ABR wave I amplitude elicited by click stimulation (Beutelmann et al., 2015) at levels between 40 and 90 dB SPL varying in 10 dB steps was measured prior to the behavioral experiment. The setup and procedure were equal to those used by Laumen et al. (2016) who measured wave I amplitudes to click stimuli in old gerbils. In short: the gerbils were anesthetized using a mixture of ketamine (70 mg/kg) and xylazine (3 mg/kg) in a 0.9% saline solution while their vital functions were monitored. A click stimulus of 33  $\mu$ s duration was played monaurally left and right in random order. The randomly varied inter-stimulus interval had an average duration of 30 ms, with a standard deviation of 10 ms. For removing possible stimulus artifacts the clicks were played with alternating signs. For each condition of level and ear at least 500 repetitions were measured and then the recordings were averaged with weights reciprocal to their signal-to-noise- ratio (Riedel et al., 2001). The Wave I peak at 90 dB SPL was visually identified. For other levels, the peaks were classified as the maximum in a  $\pm 0.45$  ms range around the 90 dB SPL peak latency with the center of this time range shifted to longer latencies by +0.2 ms per 10 dB step. The threshold was defined in a 10 dB range by the trace with the highest level where no ABR wave I peak was revealed (lower boundary) and the trace with an increased level by 10 dB revealing a wave I peak (upper boundary). For participation in the behavior experiment the gerbils needed to have at least click-evoked ABR thresholds of 60 to 70 dB SPL in each ear.

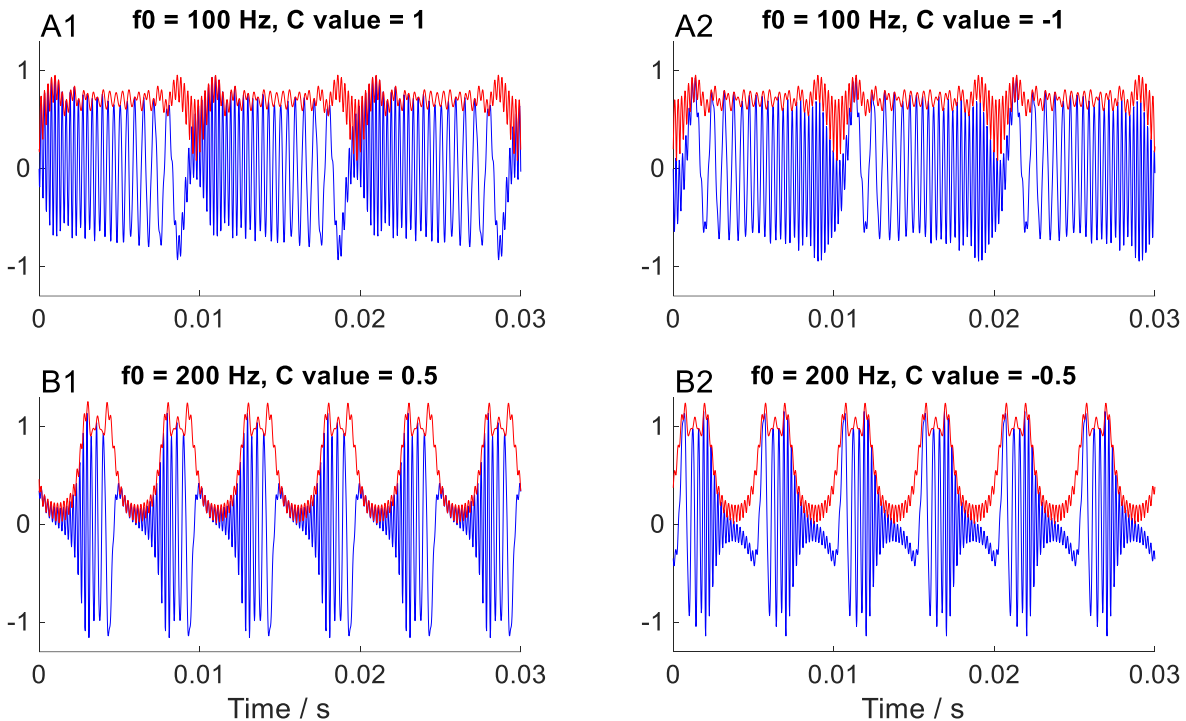
### 3.3.3 Behavioral apparatus and stimulus generation

The behavior experiments took place in a single-walled sound-attenuating chamber (Industrial Acoustics, Type 401-A). A 15 cm thick layer of sound-absorbing foam (Illbruck Illtec Pyramide 100/50, Illbruck Illtec PLANO Type 50/0) on the inside walls and on all surfaces of larger technical components ensured that the experiments were not affected by reflections and resulted in a reverberation time  $T_{60}$  for broadband noise of below 12 ms indicating a relatively anechoic environment. During the experiment, the gerbil could be observed via a closed-circuit video system under infrared wavelength illumination while in the chamber no visible light was present. In the middle of the chamber, a 30 cm long platform of fine wire mesh was installed at 90 cm height above the floor with an elevated pedestal at its center. A loudspeaker (Canton Plus XS, frequency range: 150 Hz- 21 kHz) was mounted 30 cm in front of the pedestal at 0° azimuth and elevation for a gerbil facing the loudspeaker. At the end of the platform oriented towards the loudspeaker, a food bowl was installed connected to a custom-made feeder via a tube positioned outside of the direct sound path. The gerbils were given a 10 mg custom-made food pellet as a reward for a correct response. The gerbil's position and direction on the pedestal indicating that it faced the loudspeaker was monitored by a Linux computer based on the input from a set of two custom-built light barriers. The Linux computer also signaled the feeder to provide a reward.

The stimuli were generated by custom software on a Linux-based PC and an RME soundcard (Hammerfall DSP Multiface II). The output signal was attenuated (Texio type RA-902A), amplified (Rotel type RMB 1506) and then delivered to the loudspeaker. Each day the system was calibrated with a sound-level meter (Brüel and Kjaer type 2238 Mediator) positioned on the pedestal at about the position of the gerbil to secure a maximum deviation of level between days of 1 dB SPL.

We presented Schroeder phase (SCHR) harmonic tone complexes with frequencies of equal amplitudes from the fundamental frequency up to 5 kHz. The phase of each harmonic followed this equation:  $\theta_n = C\pi n(n+1)/N$  where  $\theta_n$  was the phase of the  $n$ th harmonic,  $n$  was the harmonic number, and  $N$  was the total number of harmonics (e.g., see Leek et al., 2000). The period duration is the reciprocal of the fundamental frequency  $f_0$ . The result is a sweep within each period for which the duty cycle is determined by the absolute  $C$  value. The sweep direction is specified by the sign of  $C$  in a way that negative  $C$  values result in an increasing instantaneous frequency sweep

and positive  $C$  values in a decreasing frequency sweep within each period (Fig. 12). The discriminated stimuli within each trial had always equal absolute values of  $C$ . Thus four different pairs of  $C$  values  $\pm 0.25$ ,  $\pm 0.5$ ,  $\pm 0.75$ ,  $\pm 1.0$  had to be discriminated. Four different fundamental frequencies  $f_0$  of 50, 100, 200 and 400 Hz were chosen to account for different period durations. The overall presentation level was 60 dB SPL and the stimulus length was 0.4 s. The stimulus onset and offset were ramped by a 25 ms long Hann window. During each presentation, the starting point within the first period was randomly chosen to exclude possible onset and offset cues for discriminating. In this experiment, the gerbils discriminated a reference SCHR complex from a target SCHR complex, which only differed in the sign of the  $C$  value. Target and reference stimuli did not differ in long-term spectra or acoustical ENV. However, their acoustical TFS was time reversed which served as the only discrimination cue. We, therefore, investigated in this experiment, if gerbils can distinguish SCHR+ from SCHR- complexes only relying on acoustic TFS cues.



**Figure 12: Temporal waveforms of Schroeder phase (SCHR) complexes with different values for fundamental frequency  $f_0$  and  $C$ .**  $f_0$  determines the period, the absolute  $C$  value determines the duty cycle, the sign the sweep direction, and the combination of  $f_0$  and  $C$  value determines the rate of frequency change. The comparison of panels A1 and A2 and B1 and B2, respectively, shows the effect of the sign of the  $C$  value, i.e., the reversal of the TFS while maintaining the ENV<sub>p</sub>. Comparison of panels A1 and B1 and A2 and B2, respectively, shows the effect of halving the  $C$  value and doubling the fundamental on the duty cycle and the period of the waveform.

### 3.3.4 Procedure

In an operant conditioning Go/NoGo paradigm the gerbils were trained to wait on the elevated pedestal and face in the direction of the loudspeaker. Their position while waiting was verified by two light barriers being interrupted in the correct sequence, and the leaving of the pedestal was indicated by reconnecting at least one of the light barriers. As a background stimulus for the discrimination the reference SCHR complex was played every 1.3 s. The gerbil was trained to hop and wait on the pedestal while the background was running, which initiates a trial. After a random waiting time between 1 and 7 seconds had elapsed the next played sound was the target stimulus instead of the reference stimulus. Jumping off the pedestal before the random waiting time was reached resulted in a restart of the trial with a new random waiting time. The target trial was either a test trial with the target being the SCHR complex with

opposite-sign  $C$  value compared to the reference SCHR complex or a sham trial with the target being the same SCHR complex as the reference to assess the guessing rate. Leaving the pedestal within 1.3 s after the beginning of the test trial was registered as a “hit” and resulted in a food reward. Staying on the pedestal throughout that time period was registered as a “miss” and not rewarded. In case of a sham trial leaving the pedestal within 1.3 s after the start of the trial was registered as a “false alarm” and staying on the pedestal was registered as a “correct rejection”, and in both cases, no reward was provided. To keep the gerbils motivated and to minimize the number of false alarms a “motivation trial” with a salient test stimulus was inserted after each correct rejection and not included in the analysis.

Within a session, SCHR complexes of one sign were used as the reference and the other sign as the target. Both session types with SCHR+ and SCHR- references were alternated so that the gerbils could not get used to a specific sign of the reference  $C$  value. Each session consisted of 9 blocks with stimuli in each block having a constant fundamental frequency and varying in  $C$  value. Thus, within each session, every fundamental frequency occurred within three blocks. The first block was a warmup block with salient test stimuli and not included in the data analysis. Each block consisted of eight test trials and three sham trials. The order of blocks and the order of test and sham trials within each block were pseudo-randomized with two rules. Firstly, sham trials would follow each other within and across each block. Secondly, blocks of each of the four fundamental frequencies occurred once within the 2<sup>nd</sup> to 5<sup>th</sup> and the 6<sup>th</sup> to 9<sup>th</sup> block with the exception that the 5<sup>th</sup> and 6<sup>th</sup> block could not both contain  $f_0 = 400$  Hz stimuli. These rules were introduced to avoid within-session training effects. One additional sham trial was inserted directly after the warmup block to increase the total number of sham trials to 25.

### **3.3.5 Behavioral data Analysis**

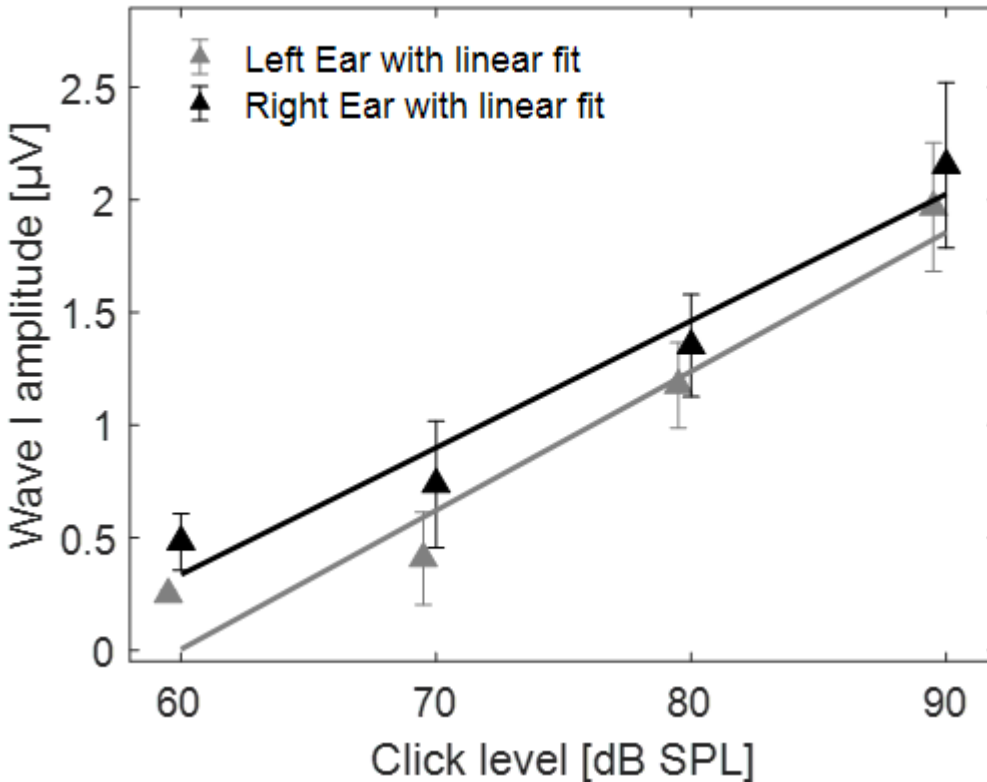
If a gerbil leaves the pedestal during the target presentation interval it could be due to the discrimination of target and reference stimuli or due to trying to get food through guessing. To distinguish between both options, only completed sessions with a maximum false alarm rate of 20% were included in the analysis. The hit rate (HR) and false alarm rate (FAR) denote the total hit and false alarm counts divided by the total number of test trials or sham trials, respectively. The discrimination sensitivity  $d' = z(\text{HR}) - z(\text{FAR})$  was derived using these rates (Green and Swets, 1966). Data collection continued by adding sessions until the sample size of each data point for a specific combination of  $f_0$  and  $C$  value was at least 20 trials. The behavioral performance and the ABRs were analyzed with general linear mixed model (GLMM) ANOVAs (SPSS 26: procedure MIXED). The significance level was 0.05 and Bonferroni correction was used for multiple comparisons.

## **3.4 Results**

### **3.4.1 Auditory brainstem responses**

The mean ABRs to click stimuli presented at levels between 40 and 90 dB SPL were measured monaurally for both ears. The thresholds of two old gerbils were 50-60 dB SPL in both ears, and the third gerbil achieved this threshold only on the left and had a threshold of 60-70 dB SPL in the right ear. A GLMM ANOVA with wave I amplitude as the dependent variable and click level and stimulated ear as factors revealed a significant increase of wave I amplitude with increasing level [ $F(3;16)=14.23$ ,  $p<0.001$ ] for the gerbils of the present study (Fig. 13). There was no significant effect of the stimulated ear or a significant interaction of both factors. In an additional GLMM ANOVA, the results of the present study were compared with data obtained from the young gerbils ( $N=7$ ) of chapter 2 using also levels between 60 and 90 dB SPL and monaural stimulation of both ears. The young gerbils had significantly higher wave I amplitudes than the old gerbils [ $F(1;64)=20.24$ ,  $p<0.001$ ] indicating age-related diminished ABRs of the old gerbils.



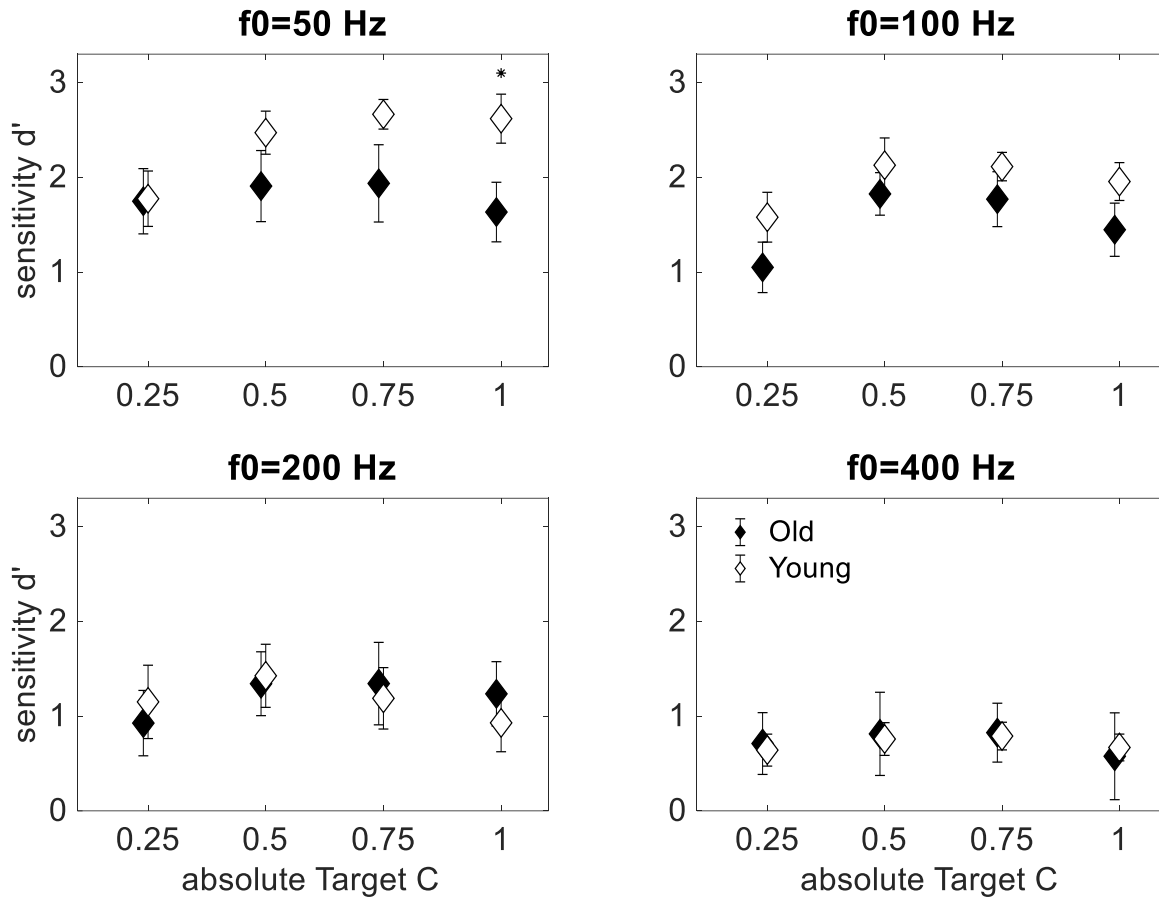


**Figure 13: ABR wave I amplitude (mean ± SE) of the three old gerbils shown for each ear separately (grey and black). The solid lines indicate the linear regressions.**

### 3.4.2 Behavioral sensitivity

We measured the behavioral sensitivity of three old gerbils for discriminating SCHR complexes only differing in the sign of  $C$ . Since there was no significant difference of sensitivity from reference SCHR+ or SCHR- complexes ( $p=0.22$ , main effect in GLMM ANOVA) the data were combined in the subsequent analysis. To illustrate the differences in sensitivity between young and old gerbils, the sensitivity (mean ± SE) is shown in Fig. 14 for different fundamental frequencies in relation to absolute target  $C$  values together with data of 7 young gerbils reported in chapter 2. The maximum achieved  $d'$  value for old gerbils was 2.25 for  $f_0 = 50$  Hz. A GLMM ANOVA with the data from old gerbils only with the sensitivity  $d'$  as the dependent variable and the factors  $f_0$  and absolute  $C$  value showed a significant main effect of  $f_0$  [ $F(3; 32)=15.05$ ,  $p<0.001$ ] and a trend for the  $C$  value [ $F(3; 32)=2.35$ ,  $p=0.091$ ] and no interaction between  $f_0$  and  $C$  value. The following pairwise comparison revealed a significant difference between every fundamental frequency, except for  $f_0 = 100$  Hz with its neighboring frequencies  $f_0 = 50$  Hz and  $f_0 = 200$  Hz, reflecting the general decrease in sensitivity with increased

fundamental frequency. Thus, SCHR complexes with  $f_0 = 50$  and 100 Hz could be much easier discriminated than with  $f_0 = 200$  and 400 Hz. A second set of GLMM ANOVAs comparing data from young gerbils (chapter 2) and old gerbils for each  $f_0$  separately with age and  $C$  value as factors and sensitivity  $d'$  as dependent variable revealed a significantly better performance of young gerbils for  $f_0 = 50$  Hz [ $F(1;32)=18.84$ ,  $p<0.001$ ] and  $f_0 = 100$  Hz [ $F(1;32)=8.84$ ,  $p=0.06$ ]. For  $f_0 = 200$  Hz and 400 Hz, the young gerbils did not perform better than old gerbils (both  $p$ -values  $>0.8$ ). No interaction was observed between  $C$  value and age in each ANOVA indicating that the effect of  $C$  value on sensitivity was similar in young and old gerbils. Overall, an age effect was only observed for conditions in which the young gerbils achieved sensitivities above a  $d'$  of 1. No further decline in sensitivity for old gerbils was measured for conditions in which already the young gerbils performed near chance.



**Figure 14: Mean sensitivity  $d'$  with SE in relation to absolute target C.** Different panels show separate data for each fundamental frequency. The young gerbils (open symbols, data from chapter 2) performed significantly better than old gerbils (filled symbols) for  $f_0$  of 50 Hz and 100 Hz. Significant pairwise comparisons between old and young gerbils within the same  $f_0$  revealed by t-tests with Bonferroni correction are marked by \*. Similarly to young gerbils, a general decrease in sensitivity with increasing fundamental frequency was observed in old gerbils.

## 3.5 Discussion

The study demonstrates that old gerbils can discriminate SCHR complexes with equal  $f_0$  and opposite signs of  $C$ . In general, with increasing  $f_0$  the sensitivity decreased, while the  $C$  value had less of an effect. A comparison with data of young gerbils from a previous study (chapter 2) revealed a significant effect of age.

### 3.5.1 Temporal perception with age

A decline in perceptual performance due to age was also observed in other behavioral studies on temporal processing in the Mongolian gerbil. Investigating temporal processing in old gerbils with little HL in a forward masking paradigm, Gleich et al. (2007) provided evidence for a slower decline in excitation in old gerbils after the masker ended. A similar age-related decline in temporal sensitivity has been observed in the form of increased gap detection thresholds of old gerbils with near-normal hearing (Hamann et al., 2004). This observation corresponds to the age-related decline in sensitivity for discriminating SCHR complexes observed in the present study. In contrast, in a vowel discrimination task (Eipert and Klump, 2020) and a difference limen task using speech continua (Sinnott and Mosqueda, 2003) no age-related decline was observed. It is possible that frequency differences between formats, i.e., spectral patterns, rather than temporal patterns provided the cues for discriminating for the vowels presented in these studies.

One could expect that an age-related decline in sensitivity will be more profound in situations not providing very salient cues. In the present study, however, the old gerbils performed only worse than the young gerbils if salient cues were present, i.e., at low  $f_0$ . No differences were observed between both age groups if the cues were not salient, i.e., at high  $f_0$ . This result is against the expectation. A similar pattern was observed in a gap detection task (Hamann et al., 2004) in which age effects only occurred at 30 dB SL providing more salient cues but not at 10 dB SL providing less salient cues. How can we explain this discrepancy from the expectation? Firstly, if there are no salient cues for young animals it is unlikely that these cues would be salient for old animals. Thus, no difference in sensitivity between old and young animals can be observed. Only if cues are above threshold for both age groups, the responses of young and old animals can be differentiated by sensitivity. Secondly, the observation that the performance between old and young gerbils differs only at low  $f_0$ s

could indicate that age-related changes at more central stages of the auditory system could be relevant. Temporal perception of SCHR complexes was also studied in humans by Lauer et al. (2009) measured discrimination of SCHR complexes with opposite  $C$  value in normal-hearing and hearing-impaired humans. The hearing-impaired humans performed worse than normal hearing humans similar to the observed effect for old vs young gerbils. Similar to the observations in the gerbil, the difference between normal-hearing and hearing-impaired humans only occurred at low  $f_0$ s but vanished with increasing  $f_0$ . This comparison with human data confirms that the gerbil is a suitable animal model to study age-related temporal processing deficits.

### **3.5.2 Physiological constraints of temporal perception**

Differences in the ongoing phase resulting in a frequency shift could provide a relevant cue for the discrimination of the SCHR complexes. Thus, a compromised ability to phase lock with increasing age could contribute to the age-related decline in temporal processing as revealed by SCHR complexes. Gleich et al. (2016) observed an almost 40 % ribbon-synapse loss at IHCs located in the apical part of the cochlea, but a relatively small reduction in the middle part of the cochlea (< 20%). Both regions are within the frequency range of the used SCHR complexes. This is emphasized by the observation that ABRs of the old gerbils in the present study were lower than those in young animals indicating peripheral HL. Comparing young and old gerbils from the same breeding facility as the gerbils of the present study, Steenken et al. (unpublished manuscript) also observed a loss of ribbon synapses that was, however, larger at the basal part of the cochlea (27 % at the 16 and 32 kHz locations) than at the apical part (12 % at the 0.5 kHz location). This result is consistent with observations by Schmiedt et al. (1996) who found an age-related ANF loss only above 6 kHz, which is above the highest component of the used SCHR complexes. The temporal coding by ANFs was not degraded in old gerbils including an unaffected phase-locking limit and phase-locking strength (Heeringa et al., 2020). Thus, there may be fewer synapses and auditory-nerve fibers in old gerbils than in young gerbils that process, however, temporal stimulus features with similar accuracy.

In general, the loss of auditory-nerve fibers, especially of low-SR fibers, has been suggested to compromise temporal coding (Bharadwaj et al., 2014; Plack et al., 2014). Despite observing age-related loss of OHCs and ANFs in humans (Wu et al., 2019), the effect on temporal perception, however, is not conclusive. In a human synaptopathy

model, the detection of interaural time differences, a temporally related task, only increases by the factor  $\sqrt{2}$  for a 50 % fiber reduction (Oxenham, 2016). Also, psychoacoustical measures found no evidence that cochlea synaptopathy in elderly humans affects temporal processing or speech reception in noise (Carcagno and Plack, 2021) and electrophysical measures did not show age-related cochlea synaptopathy in the low-frequency region of the human cochlea (Carcagno and Plack, 2020). Given that the old gerbils of the present study had reduced ABR wave I amplitudes, peripheral deficits alone are unlikely to account for the age-related decline in sensitivity for discriminating SCHR complexes. A combination of peripheral and central causes is possible since a loss or dysfunction of ANFs results in compromised signal transmission in the central system which could influence central temporal coding.

The observed decrease in sensitivity with increasing  $f_0$  in discriminating SCHR complexes suggests that more central stages in the gerbil auditory pathway may be involved. The representation of amplitude modulation and harmonic tone complexes throughout the auditory pathway has been used to describe temporal processing of periodic stimuli at different levels (Joris et al., 2004). In general, the upper limit of the representation of modulation frequencies declines with progression of the auditory pathway. At the auditory nerve, all  $f_0$ s used in the present study are well represented in most mammalian species, whereas the representation of  $f_0$ s of 200 and 400 Hz is already compromised at the level of the IC (Joris et al., 2004). Investigating responses to sinusoidally amplitude-modulated tones in the gerbil IC, Krishna and Semple (2000) observed that the best modulation frequencies in the IC were below 100 Hz for almost all tested neurons. Even a lower limit for best modulation frequencies was observed in the auditory cortex of gerbils (Schulze and Langner, 1997). The central best modulation frequencies in these physiological studies correspond to the behavioral limit of  $f_0$ , for which an age effect in SCHR perception was determined in the present study. Our result is in line with the observed compromised temporal selectivity in the IC of old gerbils (Khouri et al., 2011). Thus, central processing deficits could be a relevant factor for the age-related compromised temporal processing of SCHR complexes.

Also, developmental studies indicated a central contribution to the perception of periodic signals. Rosen et al. (2012) measured sinusoidal amplitude modulation detection threshold in normal hearing gerbils and gerbils with conductive HL both

behaviorally and neuronal with recordings from the auditory cortex. Conductive HL only affected thresholds for low modulation frequencies that were in the same order of magnitude in behavior and in cortical neurons suggesting an involvement of neurons at upper levels in the auditory pathway.

Gleich et al. (2003) and Khouri et al. (2011) suggested that perceptual deficits observed for temporal processing in old gerbils may be related to a loss of central inhibition. In a positron emission tomography study, Kessler et al. (2020) observed an age-related decrease of binding of Flumazenil to GABA receptors in the IC, medial geniculate body and auditory cortex of the gerbil suggesting a decrease of the gamma2 GABA receptor subunit with age. Also, several studies in rats found an age-related change in GABAergic neurotransmission in the IC suggesting perceptual consequences in old rodents (Caspary et al., 1990; review in 1995; Caspary and Llano, 2019). The improvement of impaired temporal processing by increasing GABAergic inhibition was experimentally demonstrated in a gap-detection task with old gerbils (Gleich et al., 2003). When treated with  $\gamma$ -vinyl-GABA (Sabril), a blocker of the GABA-degrading enzyme GABA-transaminase (Löscher and Frey, 1987), gap detection in broadband noise in old gerbils showing a perceptual deficit was improved and reached thresholds typical for young gerbils (Gleich et al., 2003). These studies suggest that a combination of peripheral HL and central deficits in GABAergic processing could explain the age-related differences of SCHR complex processing.

## **4 Chapter**

# **Discrimination of harmonic and inharmonic tone complexes based on temporal fine structure: a comparison of Mongolian gerbils with normal and compromised cochlea function**

### **Author contributions:**

Georg M. Klump designed the study. Henning Oetjen carried out the experiments. Henning Oetjen and Georg M. Klump analyzed and interpreted the data. Henning Oetjen wrote the chapter.



## 4.1 Abstract

Synaptopathy together with normal audiometric thresholds has been observed in elderly humans and old or noise-exposed animals. Therefore, synaptopathy and especially the loss of low-spontaneous rate fibers have been suggested to be a major factor for hidden hearing loss, which includes decreased sensitivity in temporal perception. A standard test for measuring monaural temporal fine structure sensitivity is the TFS1 test, in which bandlimited harmonic and frequency-shifted inharmonic tone complexes are discriminated. Here, behavioral TFS1 test discriminations for different fundamental frequencies (200 and 400 Hz) and bandpass frequency centers (1600 and 3200 Hz) were studied in the Mongolian gerbil. Age-related temporal fine structure perception was investigated by conducting the TFS1 test with untreated young (age < 18 months) and old (age > 36 months) gerbils. Additionally, young gerbils treated with a different dose of ouabain, which should lead to a loss of low-spontaneous rate fibers, participated in the TFS1 test. Measuring also the auditory brainstem responses to click stimuli in the gerbils provided an estimation of hearing loss and synaptopathy. The study demonstrated all gerbil groups were able to discriminate harmonic from inharmonic tone complexes in the TFS1 test. In general, the sensitivity decreased with increasing frequency center and decreasing fundamental frequency and frequency shift. No difference in sensitivity was observed between the untreated and the other groups of young gerbils, i.e. the treatment with ouabain did not result in lower temporal fine structure sensitivity. An age-related decline in temporal sensitivity was observed between the untreated young and old gerbils. The auditory brainstem responses indicated a loss of low-spontaneous rate fibers or synaptopathy in the ouabain treated and old gerbils. Overall, the findings suggest that peripheral deficits alone do not cause a loss of sensitivity in temporal perception. Dysfunctional central inhibition may be responsible for temporal perception deficits in old gerbils.

## 4.2 Introduction

About 10 % of adult humans, who suffer from HL have normal tone detection thresholds in quiet (Parthasarathy et al., 2020). This absence of detectable hearing deficits in the audiogram is referred to as HHL (Schaette and McAlpine, 2011). Most adults with HHL suffer from compromised supra-threshold perception in challenging environments like speech or tone in noise perception (e.g. Hind et al., 2011; Ridley et al., 2018; Ralli et al., 2019), which is often the first noticed sign of hearing deficit. Both noise exposure (e.g., Kujawa and Liberman, 2009; Valero et al., 2017) and aging (e.g., Grose and Mamo, 2010; Ruggles et al., 2012; Frisina and Frisina, 1997) can cause synaptopathy, which could lead to HHL (Wu et al., 2019).

In humans and animals, peripheral ARHL commonly is also manifested in a reduced volume of the stria vascularis (Schuknecht and Gacek, 1993; Gratton and Schulte, 1995; Ohlemiller, 2009; Schmiedt, 2010; for review: Heeringa and Köppl, 2019), which result in decreased endocochlear potentials with mostly unaffected potassium concentration in the scala media (Schulte and Schmiedt, 1992). However, a damaged stria vascularis in humans is no indicator of the degree of ARHL (Wu et al., 2020). Many studies with animal models assumed synaptopathy, i.e. the loss or dysfunction of synapses at the IHC or ANF to be a main physiological cause of HHL (e.g. Bharadwaj et al., 2014). Mammalian IHC and ANF do not regenerate once they are destroyed (Burns and Corwin, 2013; Fujioka et al., 2015). Lobarinas et al. (2013) observed only small effects on audiometric threshold for chinchillas with IHC loss up to 80%, which demonstrates processing at threshold is still possible with very few IHC. Also, 60-70 years old normal-hearing humans on average lose only 10 % of their IHC (Wu et al., 2019) and old gerbils show even a smaller loss (Tarnowski et al., 1991) so that the IHC loss supposedly cannot be the main cause for HHL. A loss of functional synapses and a degeneration of ANFs, however, could be a origin of HHL, since a loss of synaptic connection between IHC and ANF was observed due to noise exposure with only temporal and not permanent threshold shift in mice, rats and monkeys (Kujawa and Liberman, 2009; Furman et al., 2013; Valero et al., 2017). Also, old gerbils have a reduced number of ribbon synapses, mostly in the apical part of the cochlea (Gleich et al., 2016).

ANFs are categorized by their spontaneous discharge rate (SR) into low-SR, medium-SR and high-SR (Liberman, 1978; Huet et al., 2016). The low-SR fibers are most

vulnerable to noise exposure or aging (Furman et al., 2013; Schmiedt et al., 1996). These fibers have high activation thresholds and therefore do not contribute substantially to detection in quiet (Costalupes et al., 1984; Bharadwaj et al., 2014). A loss of low-SR fibers is indicated by a shallower increase of the ABR wave I amplitude with increasing level (Furman et al. 2013). The firing of medium- and high-SR fibers, however, saturates at moderate stimulus levels and is therefore not an important factor for coding stimuli in the noise.

Thus, a loss of low-SR fibers is suggested to be a major contributor to perception deficits of speech in noise and a possible physiological origin for HHL (Bharadwaj et al., 2014; Kujawa and Liberman, 2015). A method to degenerate selectively only low-SR fibers in animals is the treatment with ouabain. Bourien et al. (2014) demonstrated that mostly ribbon synapses of low-SR fibers degenerate after perfusion with ouabain at the gerbil's round window, while the numbers of medium- and high-SR fibers were not reduced. The amount of low-SR fiber loss at a given CF can be controlled by the ouabain concentration. Schomann et al. (2018) investigated the consequences of ouabain injection in the guinea pig one week after treatment and observed no loss of IHC and degeneration of OHC only for very high concentrations. They observed no loss of spiral ganglion cells, which could be due to the short time between ouabain treatment and measurement since such loss usually starts months after synaptic loss (Kujawa and Liberman, 2009; Makary et al., 2011). Batrel et al. (2017) confirmed the synapse loss caused by ouabain treatment for high frequencies. A loss of low-SR ANFs is indicated by a more shallow increase of the amplitude Wave I vs level function.

Diminished speech in noise perception reflects the inability to use TFS (the rapid oscillation), while the ENV (the slow amplitude oscillation) processing is not compromised (e.g. Lorenzi et al., 2006; Hopkins et al., 2007; Hopkins and Moore 2010). King et al. (2014) also observed higher interaural phase difference thresholds for TFS, but not for ENV differences. Thus, HHL patients should have compromised TFS perception. Henry and Heinz (2012) showed that chinchillas with SNHL due to noise exposure, as estimated from increased ABR thresholds, have degraded peripheral temporal coding in background noise. The observed broader tuning of noise-exposed fibers was made accountable for diminished temporal coding since more noise energy would enter the fiber's receptive field. Also, Hopkins and Moore (2007) demonstrated cochlea HL, indicated by air-bone gaps, leads to a reduced ability to perceive monaural TFS. A standard test for measuring monaural TFS sensitivity is the TFS1 test, which

has no learning effects and was developed by Moore and Sek (2009). Participants of the TFS1 test discriminate bandpass limited harmonic tone complexes (H) from bandpass limited inharmonic tone complexes (I), in which all frequencies are shifted upwards compared to H. The frequency shift needed for discrimination is the parameter for determining the sensitivity in TFS perception. Since auditory filters broaden with increasing level, performance-based on excitation pattern would worsen with increasing level. Moore and Sek (2011), however, observed no effect of sensation level between 10 and 50 SL for discriminating TFS1 stimuli. Also, the bandpass-filter in the TFS1 test is set to a region corresponding to unresolved harmonics. Thus, discrimination is unlikely based on excitation pattern, but rather on TFS (Jackson and Moore, 2014; Moore et al., 2009). The previously found link of diminished TFS perception (Moore et al., 2006) for patients with SNHL in comparison to normal hearing patients has also been found with the TFS1 test (Marmel et al., 2015). Furthermore, a correlation between age and TFS1 sensitivity was observed (Moore et al., 2012).

The present study investigates age-related TFS perception by conducting the TFS1 test with young and old gerbils. Additionally, gerbils treated with a different dose of ouabain, which should lead to a loss of low-SR fibers, participated in the TFS1 test. Measuring also the ABRs to click stimuli in these gerbils provides an estimation of SNHL. Comparing the TFS1 results of the gerbils with different ouabain treatments allows a deeper understanding of the influence of low-SR fibers on HHL. The present study describes the behavioral results of a joint project showing for the first time combined data of behavior, electrophysiology and histology of ouabain treated gerbils. By laying the foundation for this comparison new insight into the hypothesis that a low-SR fiber loss is a major factor for HHL is given.

## **4.3 Material and Methods**

### **4.3.1 Subjects**

Data from 24 Mongolian gerbil (*Meriones unguiculatus*) (14 male and 10 female) were obtained. They originated from Charles River Laboratories and were bred in an in-house facility of the University of Oldenburg. Five groups of gerbils were tested differing in age and ouabain treatment: named untreated ( $N=4$ ), sham surgery ( $N=4$ ), low ouabain ( $N=5$ ), high ouabain ( $N=7$ ) and old ( $N=3$ ). The first four groups comprised

young gerbils (<18 months) and the old group of gerbils older than 36 months during the time of testing

All animals were housed individually or in pairs in Type IV cages, which were provided with nesting material. Their access to water was not limited. The food given outside the experiment was limited so that the weight of the young and old gerbils was 65-75g and 70-85g respectively, to motivate the gerbils in the food rewarding measurement. All protocols and procedures were approved by the *Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES)*, Germany, permit AZ 33.19-42502-04-15/1990. All procedures were performed in compliance with the NIH Guide on Methods and Welfare Consideration in Behavioral Research with Animals (National Institute of Mental Health, 2002).

#### **4.3.2 Ouabain treatment**

The treatment with ouabain at the round window should induce a loss of low-SR fibers and is only possible via surgery in which the gerbils are anesthetized. Gerbils of the untreated group were not operated and served as a control group for normal hearing gerbils. The gerbils of the next three groups underwent surgery prior to the behavioral and ABR testing for ouabain treatment or sham treatment with physiological saline solution. These gerbils were anesthetized intra-peritoneally with a combination of ketamine (135 mg/kg; Ketamin 10%) and xylazine (6 mg/kg; Xylazin 2%) in a 0.9% saline solution. Both cochleas were exposed through a dorsal approach and a solution, which differed between the three groups, were given either manually every 10 minutes or through a syringe pump with 150 µl/h into the round window niche. After 30 min the infusion was finished. Gerbils of the sham surgery group were treated with an artificial perilymph solution (in mM: 137 NaCl, 5 KCl, 2 CaCl<sub>2</sub>, 1 MgCl<sub>2</sub>, 1NaHCO<sub>3</sub>, and 11 glucose; pH 7.4, osmolarity 304±4.3 mosmol/l) to examine the possible influence of the surgery itself. The low ouabain group was treated with the same artificial perilymph solution with 40 µM ouabain. The third operated group, the high ouabain group, was exposed to 70 µM ouabain dissolved in the artificial perilymph solution. To examine only the effects of age, the old group did not undergo any surgery.

#### **4.3.3 Auditory brainstem response recording**

Before behavioral data collection, the hearing threshold of the gerbils was evaluated. Monaural ABR evoked by click stimuli of 33 µs duration with levels ranging from 40 to 90 dB SPL in 10 dB SPL steps were recorded. The gerbils were anesthetized with an

initial mixture of ketamine (70 mg/kg), xylazine (3 mg/kg) and 0.9% saline solution. Stimulus generation and setup were described in detail in Beutelman et al. (2015). Briefly, the stimulus presentation was controlled by a custom-made software using MATLAB (The Mathworks, Inc., Natick, MA) and delivered through an RME Hammerfall DSP Multiface II sound card (RME Audio, Haimhausen, Germany), amplified by a Harman-Kardon HK6350 amplifier and broadcast by Vifa XT300/K4 speakers. A custom-built combination of exponential (100 mm length) and conical (20 mm length) horn connected the speakers with each ear. The recordings were amplified ( $\times 10,000$ ) and filtered (0.3-3 kHz) with an ISO-80 Amplifier (World Precision Instruments, Florida, USA) and sampled using the same sound card as for stimulus presentation. The sampling rate for playback and recording was 48 kHz. The randomly varied inter-stimulus interval had an average duration of 30 ms, with a standard deviation of 10 ms. For removing possible stimulus artifacts, the clicks were played with alternating signs. To secure a flat sound spectrum, the system was calibrated with an Etymotic ER-7C probe microphone positioned 2 mm into the ear canal by recording the impulse response determined prior to each measurement using a logarithmic chirp (0.2-24 kHz). The average interaural attenuation was between 30 and 40 dB below the stimulus level at each ear. The reference electrode was positioned at the vertex and the recording electrode at the neck. For each level and direction, 500 repetitions were measured and the recordings were averaged with weights reciprocal to their signal-to-noise- ratio (Riedel et al., 2001). The ABR wave I amplitudes were analyzed following the procedure used by Laumen et al. (2016). The wave I evoked by clicks at 90 dB SPL was always clearly identifiable. For other levels, peaks were classified as the maximum in a  $\pm 0.45$  ms range around the 90 dB SPL peak latency with the center of this time range shifted to longer latencies by +0.2 ms per 10 dB reduction in stimulus level. The ABR threshold was then visually identified between the maximum level for which no wave I response was observed (lower boundary) and the level increased by 10 dB (upper boundary). Except for one gerbil of the sham surgery and one of the high ouabain group, ABR thresholds of all gerbils could be identified. All gerbils, however, showed reasonable discrimination sensitivities in the behavioral training and were therefore included in the behavioral measurement. The criterion for participation in the behavioral measurement for untreated and old gerbils were thresholds of 60-70 dB SPL or better in both ears, which is the normal ABR threshold of 36 months old gerbils (Laumen et al., 2016). All measured gerbils passed this criterion, except one gerbil

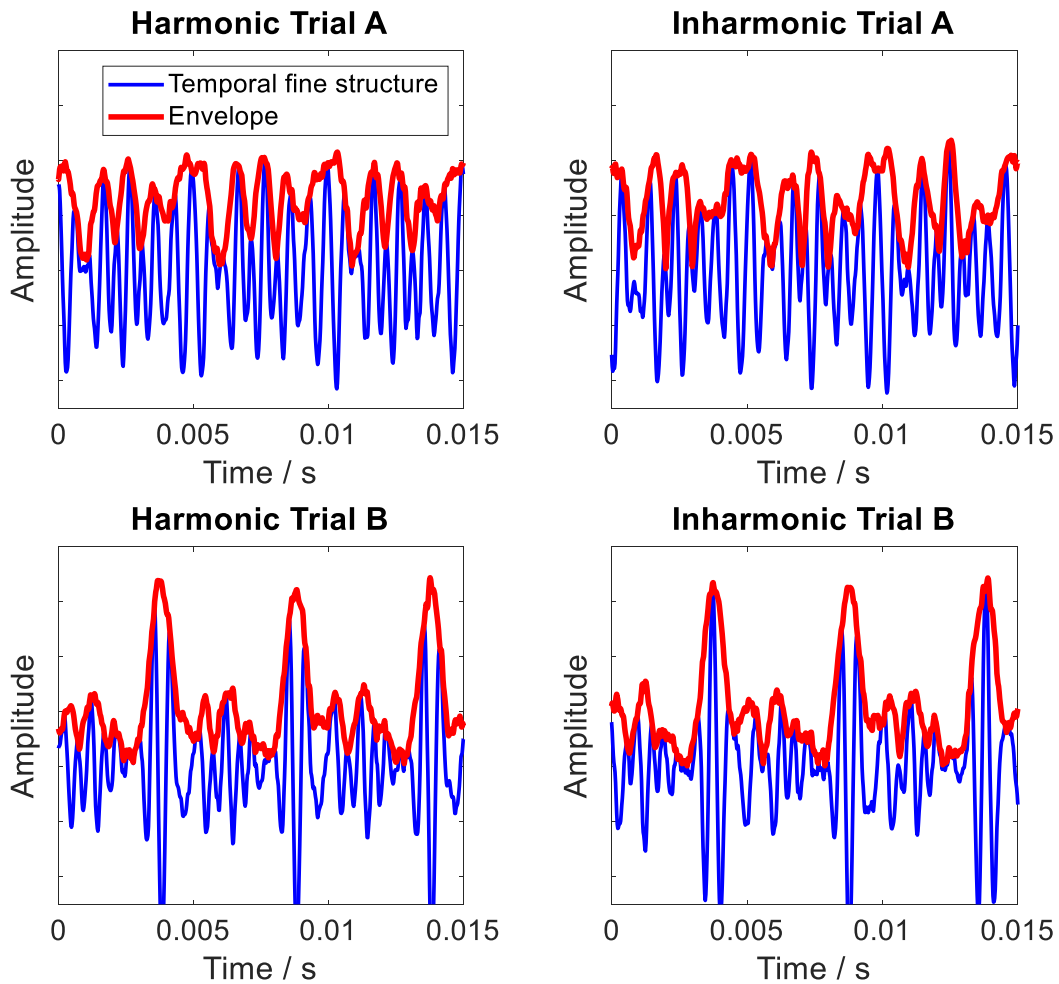
from the high ouabain group with an increased threshold in one ear (Tab. 1). Since the behavioral task requires no binaural processing that gerbil was also included in the behavioral measurement.

#### **4.3.4 Behavioral apparatus and stimuli**

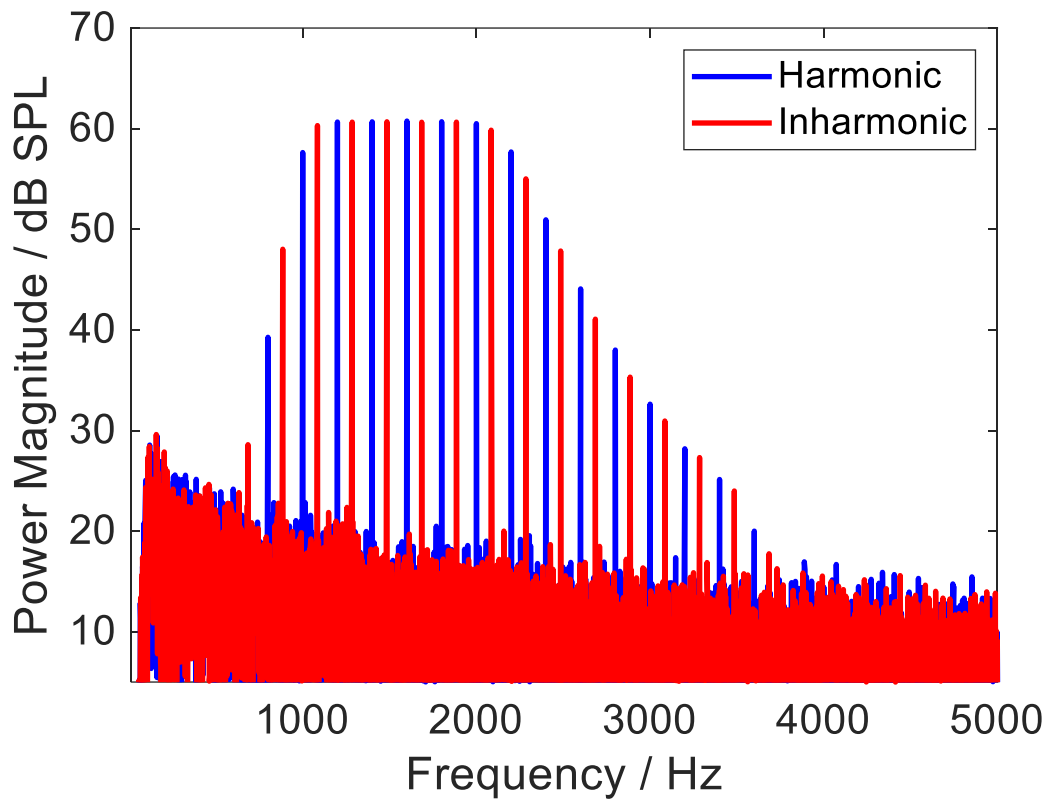
The stimuli were similar to the stimuli used for the TFS1 test in humans (for a detailed description see Moore and Sek, 2009). In short, the gerbils discriminated a harmonic tone (H) complex from an inharmonic tone (I) complex with the same  $f_0$ , but with all frequencies shifted upwards (Fig. 15). In the following, this frequency shift will be expressed in a certain percentage of  $f_0$  and will be named  $\Delta f$ . Complexes H and I were generated with components from the fundamental frequency up to 9 kHz. H and I complexes were passed through the same 5th order Butterworth infinite impulse response bandpass filter with the frequency center ( $f_c$ ) (Fig. 16) to reduce possible discrimination cues related to differences in the excitation pattern on the BM (Moore and Moore, 2003). The 3-dB down points of the filter were apart by 7 times  $f_0$  which resulted in a flat region of 5 times  $f_0$ . The filter skirts decreased by 30 dB per octave. The phase of each harmonic was randomly chosen in each trial but equal for H and I complexes in the trial to ensure similar ENV of both stimuli. A pink noise with frequencies between 100 and 11050 Hz was added to the signal to mask emerging distortion products in the inner ear and components outside the filter passband. The masker level was 20 dB SPL for the frequency of 1 kHz, which will mask distortion products of H and I stimuli in the gerbil (Faulstich and Kössl, 1999). The level of each component in the tone complex was 60 dB SPL resulting in an overall level of 68 dB SPL. The duration of each stimulus was 0.4 s. The stimulus onset and offset were ramped with a Hann window of 0.025 s. Two values of  $f_0$  (200 and 400 Hz) and two values of  $f_c$  (1600 and 3200 Hz) were used comprising three different conditions (combination of  $f_0$  and  $f_c$ ). The condition with  $f_0 = 200$  Hz and  $f_c = 3200$  Hz was not applied, because a pilot test with two young untreated gerbils showed even for very high values of  $\Delta f$  the discrimination was not possible. The harmonic rank  $N$  of a condition describes the rank of the frequency center in the harmonic tone complex. The frequency center had harmonic ranks of  $N = 8$  (condition 200/1600:  $f_0 = 200$  Hz and  $f_c = 1600$  Hz),  $N = 8$  (condition 400/3200:  $f_0 = 400$  Hz and  $f_c = 3200$  Hz) and  $N = 4$  (condition 400/1600:  $f_0 = 400$  Hz and  $f_c = 1600$  Hz). The highest value for  $\Delta f$  was 42.5 %. Other values for  $\Delta f$  were computed by dividing 42.5 % by  $2^n$ , with  $n$  being 1 or

increasing integers. The lowest measured  $\Delta f$  was 0.33 %. Additionally, the values of  $\Delta f = 31.87\%$  and  $15.94\%$  were used to ensure the old gerbils were presented with a sufficient number of trials allowing a salient percept for discrimination. The periodic ENV of H and I complexes are similar and offer no cue for discrimination while the TFS of H and I complexes are not identical. The TFS of the I complex changes across each period (most obvious for the TFS in the peaks of the ENV in Fig. 15) while the TFS of the H complex does not. Only this difference can be used for discrimination. The small random change in the TFS and ENV of both H and I complexes due to the addition of the masking noise offer no discrimination cue. One concern with the TFS task is that there might be excitation pattern cues, i.e. excitation of different auditory filters by harmonic and inharmonic tone complexes. To assess the availability of such cues, the filter bandwidths in the gerbil were calculated using the critical-band function (Greenwood, 1961) and parameters derived from critical ratio measurements (Kittel et al., 2002). No difference in excitation pattern between harmonic and inharmonic tone complexes was observed for each condition and thus also the gerbils could only use TFS cues for discrimination.





**Figure 15: Temporal waveforms of harmonic (H) and inharmonic (I) tone complexes in the TFS1 test** for two different trials A and B with TFS in blue and ENV in red for condition 200Hz/1600Hz with  $\Delta f = 42.5\%$ . The ENV of H and I complexes are similar and offer no cue for discrimination. The TFS of the complex I changes across each period, while the TFS of the complex H does not. This difference in TFS offers the cue for discrimination in the TFS1 test.



**Figure 16: Power spectra of H (blue) and I (red) complexes with frequency shift  $\Delta f = 42.5\%$  of  $f_0$  embedded in pink noise.** The analysis bandwidth was 1 Hz. The complexes had an  $f_0$  of 200 Hz and a center frequency of 1600 Hz. The pink-noise masker had an upper-frequency limit of 11050 Hz. The masking noise was at 1 kHz 40 dB below the harmonics.

The behavioral experiments were conducted in a single-walled sound-attenuating chamber (Industrial Acoustics, Type 401 A) lined with a 15 cm thick layer of sound-absorbing foam (Illbruck Illtec Pyramide 100/50 on Illbruck Illtec PLANO Type 50/0). All devices within the chamber produced no relevant sound reflections. The time for the reverberation in the chamber to decay by 60 dB was 12 ms indicating nearly anechoic conditions.

In the center of the chamber, a 30 cm long platform was located at 90 cm height with an elevated pedestal on its center, where the gerbils waited during the measurement. This construction was made of wire mesh to limit sound reflections. At end of the platform, a food bowl was located attached via a tube outside of the sound path to a custom-made feeder rewarding the gerbils for correct discrimination with a 10-mg custom-made food pellet. A loudspeaker (Canton Plus XS, frequency range: 150 Hz - 21 kHz) was mounted 30 cm in front of the pedestal at 0° elevation and azimuth relative to the gerbil's normal head position. Out of the gerbil's reach, a system of two custom-made light barriers was installed to detect the gerbil's position and facing direction on the pedestal. An infrared camera was positioned above the platform to observe the gerbil's movement on the platform while the measurement was conducted under invisible infrared light. Outside the chamber, the camera image was displayed on a monitor. The stimulus presentation, registration of light barriers switching and the feeder were controlled by custom-written software on a Linux-based PC with an RME sound-card (Hammerfall DSP Multiface II, sampling frequency 48 kHz). The output signal from the sound card was manually attenuated (Texio type RA-902A), amplified (Rotel type RMB 1506) and presented by the loudspeaker. Before each measurement day, the system was calibrated with a sound level meter (Brüel and Kjaer type 2238 Mediator) positioned on the elevated pedestal to secure a maximum deviation of level between days of  $\pm 1$  dB.

#### **4.3.5 Procedure**

The behavioral experiment used an operant Go/NoGo paradigm with food reward. Gerbils were trained to wait on the elevated pedestal and face in the direction of the loudspeaker. During the whole session, the H complex was played every 1.3 s as a reference stimulus. If the gerbil jumped onto the pedestal and interrupted the light barriers in the correct sequence, a trial was started and a random waiting time between 2 and 7 seconds was chosen. Waiting on the pedestal the complete waiting

time indicated by continuously interrupting both light barriers resulted in playing a target stimulus instead of the reference H complex. Leaving the pedestal before the waiting time elapsed resulted in the restart of the trial with a new waiting time. The target stimuli could either be equal to the reference H complex (sham trial) for testing the guessing rate or different (test trial: I complex) for testing the discrimination ability. Leaving the pedestal within 1 s after the onset of the target stimulus of a test trial, as was indicated by the light barriers, was registered as a “hit” and resulted in a food reward. Staying on the pedestal in a test trial was counted as a “miss” with no food reward. The animal had to leave the pedestal and jump on it again in the correct sequence to start a new trial after a “miss” occurred. Leaving the pedestal within 1.3 s after the start of the target stimulus during a sham trial was registered as a “false alarm” while staying on the pedestal in a sham trial was registered as a “correct rejection”. No food reward was provided in sham trials. To keep the gerbils motivated during the session and to minimize the number of false alarms a salient test trail not included in the analysis was inserted after each correct rejection.

Each session contained 10 blocks with 9 trials each (six test trails, three sham trials). All stimuli within a given block had the same condition of either 200/1600, 400/1600 or 400/3200. The first block was a warmup block with all test trials shifted by  $\Delta f = 42.5\%$  and was not included in the analysis. To interleave simple and more difficult conditions, the subsequent blocks of different conditions were pseudo-randomly distributed across the session. The least salient condition 2/16 did not occur within two consecutive blocks. Additionally, the sham and test trials within each block were pseudo-randomly distributed with the constraint that no sham trials followed one another within and across blocks. Initial training of the gerbils started with the expected most salient condition 4/16 and  $\Delta f = 42.5\%$  to familiarize the gerbils with the waiting on the pedestal and the discrimination of the acoustic stimuli. Next, the other conditions were added and in the last step, the five lower  $\Delta f$  steps (31.87; 21.25; 15.94; 10.63; 5.31 %) were introduced. The training ended if the gerbil achieved a false alarm rate below 20 % in three consecutive sessions with reasonable high hit rates for the three highest  $\Delta f$  values.

In testing, sessions were included in the analysis if the gerbil completed all 90 trials and the false alarm rate did not exceed 20%. Collecting data were continued until the sample size for each condition and data point was at least 20 trials. The sensitivity  $d' =$

$z(\text{HR}) - z(\text{FAR})$  was determined by applying the z-transform to the hit rate (HR) and the false alarm rate (FAR). Not all gerbils were in all conditions equally sensitive resulting sometimes in  $d'$  values above 1 for all initial six  $\Delta f$  values. Since the threshold of  $d' = 1$  should be determined,  $\Delta f$  was adjusted after the sensitivity with the initial six  $\Delta f$  values were measured. Therefore, lower  $\Delta f$  values were added into the set of six  $\Delta f$  values with higher more salient  $\Delta f$  values were cut out. The measurement was repeated with the new set of  $\Delta f$  values until again the sample size of each data point was at least 20. The sensitivity decreases with decreasing  $\Delta f$ , but the six starting  $\Delta f$  steps did not result for each gerbil in sensitivities below the threshold of 1. The sensitivity decrease, however, allowed an estimation, for each condition separately, for which  $\Delta f$  the sensitivity would fall below the threshold of  $d' = 1$ . Based on this estimation new lower  $\Delta f$  values were added into the set of six  $\Delta f$  values and higher more salient were cut out so that the lowest  $\Delta f$  would presumably result in  $d' < 1$ . The measurement was repeated with the new set of  $\Delta f$  until again the sample size of each data point was at least 20. This procedure was repeated until the measured sensitivity of at least one data point was below  $d' = 1$  for each condition.

#### **4.3.6 Statistical Analysis**

The behavioral sensitivities and the ABR wave I amplitudes were analyzed with general linear mixed model (GLMM) ANOVAs (SPSS 26: procedure MIXED). The significance level was 0.05 and Bonferroni correction was used for multiple comparisons. As common in signal detection theory, the threshold for discrimination in the behavioral experiment was set to  $d' = 1$ .

## **4.4 Results**

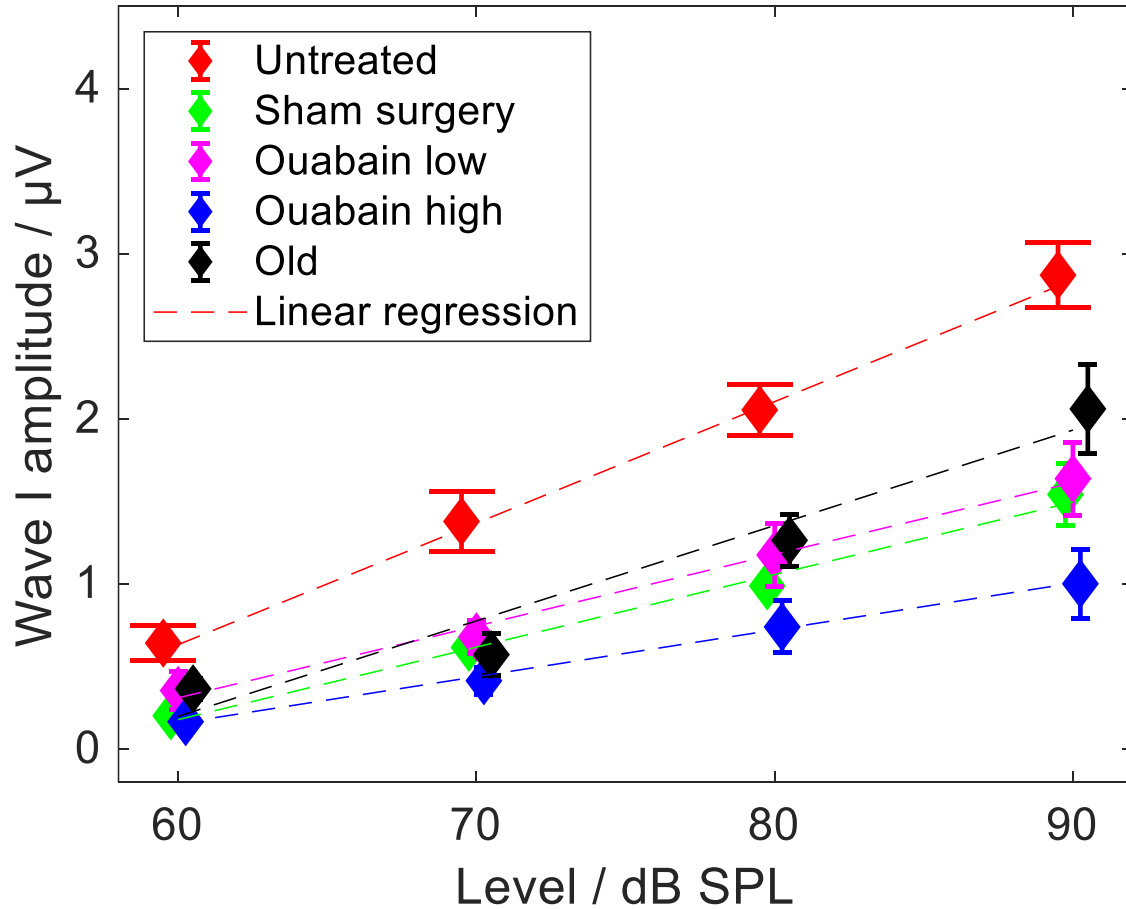
### **4.4.1 Auditory brainstem response**

The mean ABRs to click stimuli were measured monaurally for both ears at levels between 40 and 90 dB SPL. Most of the gerbils from the untreated, sham surgery, low ouabain and old group had thresholds for both ears of 50-60 dB SPL (Tab. 1). These visually derived thresholds are similar to the 50-60 dB SPL thresholds of young gerbils in a study from Laumen et al. (2016) but lower than the 60-70 dB SPL thresholds of most old gerbils from the same study. The on average 10 dB increased thresholds for most high ouabain gerbils are similar to the thresholds of most old gerbils from

Laumen et al. (2016). A GLMM ANOVA with wave I amplitude as dependent variable and click level, stimulated ear and group as factors revealed a significant effect for stimulated ear [ $F(1,133)=12.69$ ,  $p=0.001$ ]. Since possibly the TFS coding of the better ear is interrupted by the weaker ear in the monaural TFS1 test the mean between both ears is shown in Fig. 17. Furthermore, the ANOVA revealed a significant effect of level [ $F(3,133)=76.70$ ,  $p<0.001$ ], group [ $F(4,133)=33.84$ ,  $p<0.001$ ] on wave I amplitude, and the interaction between both factors [ $F(12,133)=2.47$ ,  $p=0.006$ ] was also significant. The pairwise comparison revealed a significantly increased wave I amplitude for each level increase (all  $p<0.001$ ). The untreated group achieved larger wave I amplitudes than all other groups (all  $p<0.001$ ), while the amplitudes of the high ouabain group were lower than that of all other groups (all  $p\leq 0.016$ ).

**Table 1: ABR threshold in response to click stimuli in dB SPL for left and right ear with the indication of group, age and treatment for all tested gerbils.** The threshold was measured in 10 dB steps and visually identified between the maximum level for which no wave I response was observed (lower boundary) and the level increased by 10 dB (upper boundary).

Group	Age/month	Treatment	Left	Right
Untreated	<18	None	50-60	40-50
Untreated	<18	None	50-60	40-50
Untreated	<18	None	50-60	50-60
Untreated	<18	None	60-70	50-60
Sham surgery	<18	Sham solution	50-60	50-60
Sham surgery	<18	Sham solution	40-50	50-60
Sham surgery	<18	Sham solution	50-60	50-60
low Ouabain	<18	40 $\mu$ M Ouabain	40-50	40-50
low Ouabain	<18	40 $\mu$ M Ouabain	60-70	60-70
low Ouabain	<18	40 $\mu$ M Ouabain	50-60	40-50
low Ouabain	<18	40 $\mu$ M Ouabain	60-70	50-60
low Ouabain	<18	40 $\mu$ M Ouabain	40-50	40-50
high Ouabain	<18	70 $\mu$ M Ouabain	60-70	60-70
high Ouabain	<18	70 $\mu$ M Ouabain	50-60	70-80
high Ouabain	<18	70 $\mu$ M Ouabain	60-70	60-70
high Ouabain	<18	70 $\mu$ M Ouabain	60-70	60-70
high Ouabain	<18	70 $\mu$ M Ouabain	60-70	60-70
high Ouabain	<18	70 $\mu$ M Ouabain	60-70	50-60
high Ouabain	<18	70 $\mu$ M Ouabain	60-70	50-60
Old	>36	None	50-60	50-60
Old	>36	None	50-60	50-60
Old	>36	None	60-70	50-60



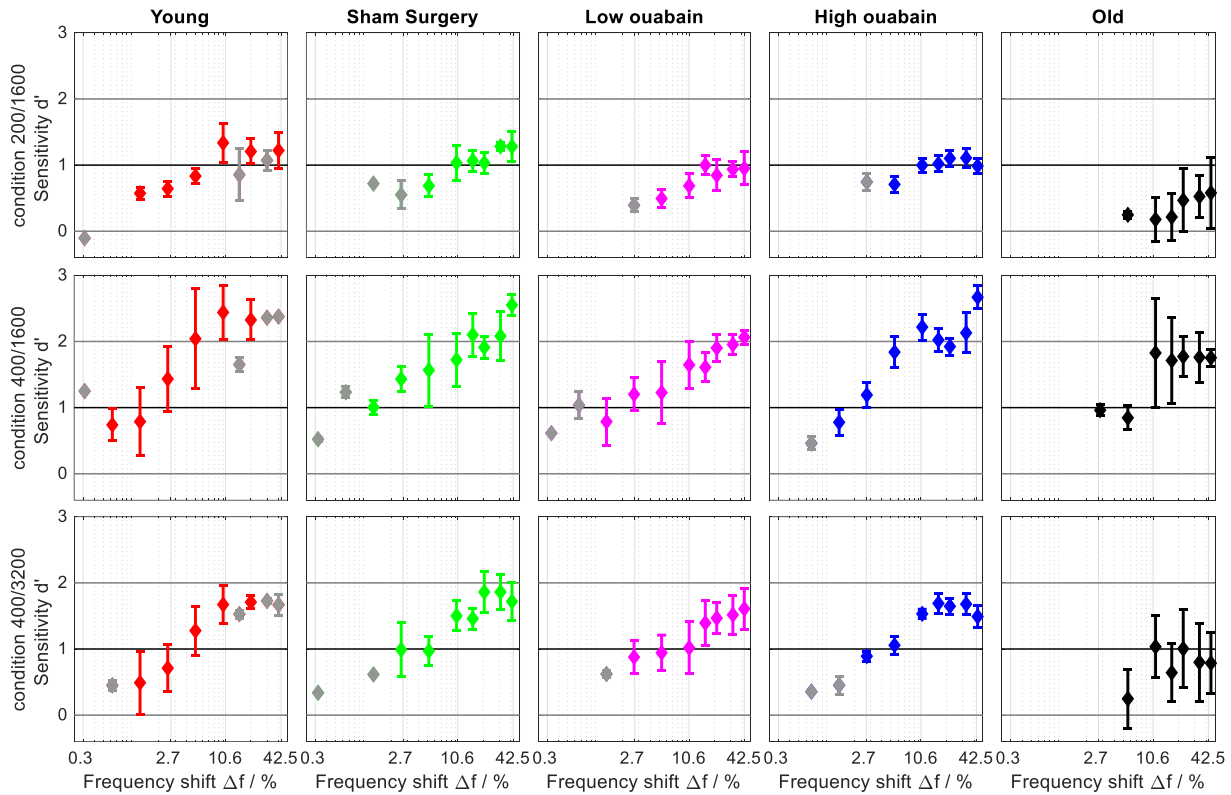
**Figure 17: Effect of click level in dB SPL on wave I amplitude (mean  $\pm$  SE) in  $\mu\text{V}$  for both ears combined.** The color indicates the different gerbil groups and the dashed lines the linear regressions. The untreated gerbils showed the largest and the high ouabain group the lowest wave I amplitudes compared to all other groups.

#### 4.4.2 Behavior sensitivity

The behavioral sensitivity was measured from 21 young gerbils of the untreated group ( $N=4$ ), the sham surgery group ( $N=4$ ), the low ouabain group ( $N=5$ ) and the high ouabain group ( $N=8$ ) together with the old group ( $N=3$ ) for discriminating harmonic and inharmonic tone complexes based on TFS. To illustrate the differences in sensitivity between the five groups, the mean sensitivity with SE in relation to  $\Delta f$  is shown for each measured condition, 400/3200, 200/1600 and 400/1600 in separate panels (Fig. 18A-C). A GLMM ANOVA with the sensitivity  $d'$  as dependent variable and condition,  $\Delta f$  and treatment group as factors revealed significant main effects of condition [ $F(2, 382)=70.45$ ,  $p<0.001$ ],  $\Delta f$  [ $F(10, 382)=14.37$ ,  $p<0.001$ ] and group [ $F(4, 382)=13.22$ ,  $p<0.001$ ]. No interaction was observed between any of the factors. The pairwise



comparison revealed a general decrease in sensitivity with decreasing  $\Delta f$ . It was also revealed that old gerbils showed lower sensitivities than all young gerbil groups (all  $p \leq 0.009$ ). The higher sensitivity of the high-ouabain group than the low-ouabain group marks the only difference within the groups of young gerbils ( $p = 0.047$ ). Investigating the effects of  $f_0$  and  $N$  on sensitivity is equivalent since both parameters increase by the same factor when comparing the conditions 400/1600 and 400/3200. To evaluate the effects of  $f_0$  and  $f_c$  respectively  $N$  a planned comparison was conducted in two additional GLMM ANOVAs, each containing only the data of two conditions, with the same values as dependent variable and factors as the initial ANOVA. The planned comparison with the data of conditions 400/3200 and 200/1600 with equal harmonic rank  $N = 8$  (Fig. 18A and B) is used to investigate the effect of  $f_0$  on sensitivity. This ANOVA revealed an effect of  $f_0$  with a higher sensitivity for increasing  $f_0$  [ $F(1, 244) = 30.42, p < 0.001$ ]. Furthermore, an effect of  $\Delta f$  [ $F(10, 244) = 9.19, p < 0.001$ ] was observed. The main effect of group [ $F(4, 244) = 16.92, p < 0.001$ ] occurred only due to differences between the old group and all other groups (all  $p < 0.001$ ). The second planned comparison with the data of conditions 400/3200 and 400/1600 with equal  $f_0$  (Fig. 18A and C) investigates the dependency of  $f_c$  respectively  $N$  on sensitivity. An effect of  $f_c$  respectively  $N$  [ $F(1, 265) = 36.33, p < 0.001$ ] was shown with decreasing sensitivity for increasing  $f_c$  respectively  $N$ . Also this ANOVA revealed decreased sensitivity with decreasing  $\Delta f$  [ $F(9, 265) = 14.40, p < 0.001$ ], while the main effects of group [ $F(4, 265) = 7.73, p < 0.001$ ] was only due to differences between the old and the untreated gerbils ( $p = 0.008$ ). No interaction was observed in both planned comparisons.



**Figure 18: Sensitivity  $d'$  (mean  $\pm$  SE) in relation to frequency shift  $\Delta f$  between the harmonic and inharmonic complex.** Panels show different conditions (vertical direction) and groups (horizontal direction). The color scheme is equal to Fig. 17. Data points measured with less or equal than half of the gerbils within each group are indicated in grey. Black lines indicate the threshold of  $d' = 1$  for visual guidance.

To derive the frequency shift  $\Delta f$  needed for the threshold sensitivity of  $d' = 1$  (a common threshold in signal detection theory), the data points directly below and above  $d' = 1$  were linearly interpolated for each gerbil and condition. No threshold could be derived if all measured sensitivities for one gerbil and condition were below 1. These values were excluded from the mean value calculation (Tab. 2). The old gerbils were only able to discriminate the TFS1 stimuli for the condition 400/1600, but not for the other conditions. All young gerbil groups achieved for each condition sensitivities above threshold, except both ouabain groups for the condition 200/1600. Next, a GLMM ANOVA was conducted with rank transformed  $\Delta f$  thresholds as dependent variable and condition and group as factors. For this purpose, the values for which no threshold could be calculated were set to the maximum rank +1. Similar to the ANOVA with all  $\Delta f$  values included, main effects for condition [ $F(2, 57)=9.71, p<0.001$ ] and treatment

group [ $F(4, 57)=3.26, p<0.021$ ] on  $\Delta f$  threshold were revealed with no interaction between both factors. The pairwise comparison revealed old gerbils had higher thresholds than gerbils of the untreated group ( $p=0.02$ ) and the sham surgery group ( $p=0.04$ ). No differences were observed within the young gerbil groups. To evaluate the effect of  $f_0$  on the threshold a planned comparison with the data of conditions 400/3200 and 200/1600 was conducted in an additional ANOVA with rank transformed thresholds as dependent variable and condition together with gerbil group as factors. No main effect for condition and gerbil group was revealed, indicating  $f_0$  did not affect the threshold. A second additional ANOVA with the data of conditions 400/3200 and 400/1600 investigates the dependency of  $f_c$  respectively  $N$  on the threshold. An effect of gerbil group [ $F(4, 38)=2.73, p<0.043$ ] and condition [ $F(1, 38)=6.29, p<0.017$ ] was observed. The pairwise comparison revealed lower thresholds with decreasing  $f_c$  respectively  $N$  and a significant difference between untreated and old gerbils ( $p=0.05$ ). No interaction was observed in both additional planned comparisons.

**Table 2: Mean  $\Delta f$  threshold ( $d' = 1$ ) in relation to condition and group.** No threshold was derived and marked as x if half or more gerbils within a particular condition and group did not achieve sensitivities above the threshold of  $d' = 1$ .

Group	Condition	Condition	Condition
	400/3200	200/1600	400/1600
Untreated	5.03	7.59	2.91
Sham surgery	6.49	5.52	4.44
Low ouabain	15.47	x	6.86
High ouabain	4.50	x	2.47
Old	x	x	10.74

## 4.5 Discussion

The study demonstrated that gerbils of the untreated, sham surgery, low ouabain, high ouabain and the old group can discriminate harmonic from inharmonic tone complexes in the TFS1 test. No difference in sensitivity was observed between the untreated and the other groups of young gerbils, i.e. the treatment with ouabain did not result in lower TFS sensitivity. The old gerbils performed worse than the untreated gerbils in both measures, discrimination thresholds and sensitivities with all  $\Delta f$  values included. In general, the sensitivity decreased with increasing  $f_c$  respectively harmonic rank  $N$  and decreasing  $f_0$  and  $\Delta f$ . The thresholds decreased with decreasing  $f_c$  respectively  $N$  but were unaffected by a change in  $f_0$ . In the ABR measurement, the untreated group showed higher wave I amplitudes, than all other groups.

### 4.5.1 TFS1 sensitivity in young and old subjects: comparison between humans and gerbils

Sensitivity to TFS has also been studied with the TFS1 test in humans before. Many studies with young normal-hearing humans showed, similar to the young gerbils, a decrease in sensitivity in the TFS1 test with increasing  $f_c$  respectively  $N$  and equal  $f_0$  (e.g. Moore and Sek, 2009; Moore et al., 2009; Hopkins and Moore, 2011; Jackson and Moore, 2014). In predecessor studies of the TFS1 test for investigating pitch perception in humans also a reduced benefit from TFS with increasing harmonic rank in a complex tone was observed (Moore and Moore, 2003a and b).

They proposed that a reduction in the precision with which inter-spike intervals in a given auditory neuron can be measured could be responsible for the decreased sensitivity at high  $f_c$ , similar to the compromised phase-locking at high CF. The sensitivity in humans also decreases with decreasing  $f_0$  and constant  $N$ , however only for  $f_0 \leq 100$  Hz. Human sensitivity shows generally high values between  $f_0 = 100$  to 400 Hz (Moore and Sek, 2009; Hopkins and Moore, 2010; Hopkins and Moore, 2011; Jackson and Moore, 2014; Füllgrabe and Moore, 2015) and only decrease again for very high  $f_0$  up to 800 Hz (Moore and Sek, 2009b). Thus, decreasing sensitivity with decreasing  $f_0$  from 400 Hz to 200 Hz, like in the gerbil, is not observed in humans. Discrimination of TFS1 stimuli on a neural level has been suggested to be possible by evaluating the time period between peaks in the TFS within each auditory filter since the time period between these peaks is lower for inharmonic than for harmonic tone

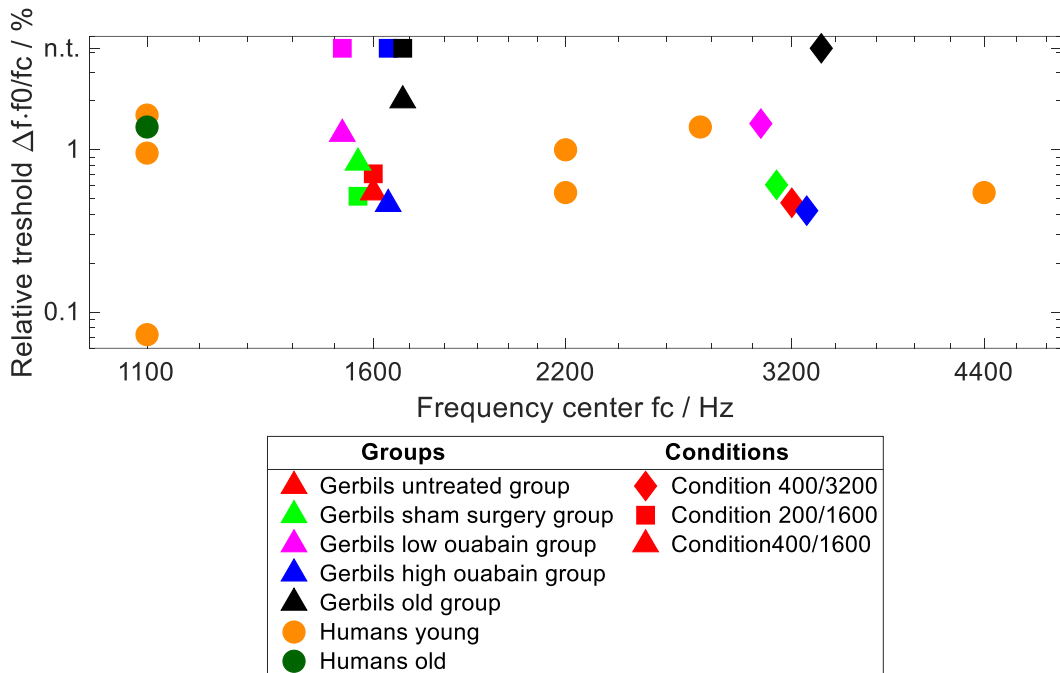
complexes (Moore et al., 2014). The decrease in sensitivity for low  $f_0$  is suggested to occur because the auditory system has difficulties in measuring very long inter-spike intervals (de Cheveigne and Pressnitzer, 2006) or because relatively few periods of the sound occur within the integration time of the ear (Moore and Sek, 2009). A main effect for  $f_0$  is revealed by comparing conditions with equal harmonic numbers and therefore different  $f_c$ . The gerbil has fewer ANFs, especially low-SR fibers, in the frequency region of the TFS1 stimuli (Schmiedt, 1989), which may deviate from the ANF distribution in humans. Therefore, the number of ANFs could also be responsible for differences in sensitivity between both species for  $f_0 = 200$  and 400 Hz.

An age-related decline in temporal perception in the Mongolian gerbil was also observed in other temporal tasks like gap detection (e.g. Hamann et al., 2004), forward masking (e.g. Gleich et al., 2007) and in the SCHR complex discrimination (chapter 2). Thus, the age-related decline of sensitivity in the TFS1 test is in line with previously reported compromised temporal processing of old gerbils. The observed decreased wave I amplitudes of the old gerbils in comparison to the untreated young group suggests that synaptopathy could be at least partly responsible for the age-related decline. However, temporal coding of the remaining ANFs in old gerbils is not degraded (Heeringa et al., 2020). Thus, not the decreased precision in the temporal coding of single ANFs, but only the quantitative loss of ANFs in old gerbils could account for reduced temporal sensitivity (Bharadwaj et al., 2014). Furthermore, a stria vascularis dysfunction in combination with ANF loss, as has been commonly observed due to aging in many species (Heeringa and Köppl, 2019), could also have negatively affected temporal perception of the old gerbils. However, in elderly humans, the degree of HL was predicted by the loss of hair cells and ANFs and not by stria vascularis dysfunction (Wu et al., 2020).

Similar to the results obtained with gerbils, several studies in humans measured an age-effect on sensitivity in the TFS1 test, although the effect could not always be separated from an effect of HL (e.g. Hopkins and Moore, 2010, 2011). Moore et al. (2012) measured TFS1 sensitivity for normal-hearing humans varying in age and observed a correlation between age and TFS1 sensitivity. This effect was also confirmed in other studies applying the TFS1 test in young and elderly normal-hearing listeners (Füllgrabe et al., 2015; Eipert et al., 2019). Age-related deficits in temporal processing

have also been found for humans in other temporal tasks like SCHR complex discrimination (e.g. Lauer et al., 2009), gap detection (e.g. Snell 1997) and forward masking (e.g. Strouse et al. 1998). Thus, the observed age-related decline of performance in temporal processing in the gerbil is in line with previously found compromised temporal processing of elderly humans.

To compare absolute TFS1 thresholds between gerbils and humans the relative threshold ( $\Delta f \cdot f_0 / f_c$ ) is shown together with data of young normal-hearing and elderly humans from Moore and Sek (2009); Jackson and Moore (2014); Marmel et al. (2015) and Eipert et al., 2019 (Fig. 19). For humans, all harmonics were unresolved since  $f_c$  was centered at  $11 \cdot f_0$ . The human data were obtained in a 2 alternative-forced-choice procedure, in which the 70.7% correct response threshold corresponds to the  $d' = 1$  threshold in the Go/NoGo procedure used for the gerbil experiment. Overall, the thresholds of the untreated group of young gerbils are for each condition in the range of the thresholds of young humans, with the human data point for  $f_c = 1100$  Hz and a relative threshold of 0.08 being an outlier. Old gerbils have higher thresholds than young and old humans.



**Figure 19: Comparison of relative TFS1 thresholds ( $\Delta f \cdot f_0 / f_c$ ,  $\Delta f$  thresholds from Tab. 2) for gerbils of the different groups and conditions with performance of humans as a function of frequency center  $f_c$ .** The color indicates the gerbil group and humans of different ages. No threshold measurable is indicated with n.t. (all  $d'$  below 1) and the symbols indicate the different conditions for the gerbils (diamond, square and triangle) Frequency center  $f_c$  was always  $11 \cdot f_0$  with  $f_0$  ranging from 100 Hz to 400 Hz for the human data (Young: Moore and Sek (2009); Jackson and Moore (2014); Marmel et al. (2015); Young and old: Eipert et al., 2019). For both, gerbils and humans, all harmonics were unresolved.

#### 4.5.2 Effects of synaptopathy on temporal perception

The experiments on the ouabain-treated groups aimed to investigate the effects of a low-SR fiber loss on temporal perception. However, both ouabain-treated groups showed no reduced sensitivity in the TFS1 test in comparison to the untreated group. Mathew et al. (2016) measured the performance of young normal-hearing subjects and subjects with SNHL, which could include a loss of low-SR fibers, being on average only 8 years older in the TFS1 test and observed an effect of HL on temporary processing. In a meta-study combining a large data set of the binaural TFS-LF test, Füllgrabe and Moore (2018) separately examined the influence of audiometric threshold and age on TFS perception. They confirmed the reduced binaural TFS sensitivity with HL and found a stronger performance reduction due to aging than due to auditory threshold

elevation. Thus, the lack of aging effect of the ouabain treatment that should have produced sensorineural damage does not correspond to the findings in humans, although also in the gerbil the effect of age prevails. Comparing absolute values, gerbils of both ouabain groups had similar relative thresholds than humans for the conditions 400/1600 and 400/3200, while no thresholds could be measured in the gerbil for the condition 200/1600 (Fig. 19). Furthermore, suggesting the number of synapses is not a good predictor for TFS1 sensitivity. The surgery itself had no influence on perception in the TFS1 test since the sham surgery group did not show compromised behavioral performance at threshold and for all  $\Delta f$  included.

Synaptopathy is suggested to be a major factor for temporal processing deficits related to HL (Lorenzi et al., 2006). In line with that, the ABR data showed higher audiometric thresholds for the high ouabain group and lower slopes of ABR growth function with increasing intensity for both ouabain groups. However, the uncompromised performance of both ouabain groups in the behavioral TFS1 test suggests that a loss of low-SR fibers has no effect on TFS perception in the gerbil. According to Bourien and colleagues (2014), gerbils of the low ouabain group lose almost all low-SR fibers related to frequencies below 4 kHz, while the high ouabain group also lose medium and high SR-fibers in that frequency range. Furthermore, both ouabain-treated groups should lose more fibers in total in the high-frequency range. They proposed that the high jitter in the timing of spikes in conjunction with the long first spike latencies of low-SR fibers (e.g. Huet et al., 2016) make the low-SR fibers unlikely to contribute to CAPs. Furthermore, at levels near audiometric thresholds mainly the medium and high-SR fibers are firing (Liberman, 1978; Bharadwaj et al., 2014). Thus, the low-SR fibers likely contribute less to the ABR at the audiometric threshold than the other fiber types which explains the higher ABR thresholds of the high ouabain group and the lack of threshold shift for the low ouabain group. Similar to the low ouabain group, reduced suprathreshold amplitudes in absence of ABR threshold shifts were also observed in guinea pigs, which lost due to noise exposure up to 30% of the entire ANF population while the properties of the remaining ANFs were intact (Furman et al., 2013). This loss of ANFs was attributed to low-SR fibers. A loss of ANFs was also revealed in elderly normal-hearing humans (Wu et al., 2019).

Although the ABR data suggest a synaptopathy in the ouabain-treated gerbil, they showed no reduced sensitivities in the TFS1 test. There may be several explanations



for this observation. It is possible that not a sufficient number of low-SR fibers in the frequency region below 5 kHz were degenerated by the ouabain treatment to compromise sensitivity in the TFS1 test. Batrel et al. (2017) also only observed a synapse loss in the 16 and 32 kHz regions for gerbils treated with 33  $\mu$ M ouabain, not differing between SR. The infusion at the round window naturally affects the fibers located in the high-frequency region of the cochlea more. In addition, the low-SR fibers in the gerbil are not evenly distributed over all frequencies but are mainly located in the range  $>4$  kHz (Schmiedt, 1989; Huet et al. 2016). Furthermore, the assumption that medium and high-SR fibers saturate at supra-threshold levels while low-SR fibers still represent the stimulus is only valid at the CF (Liberman, 1978). The off-CF medium and high-SR fibers could compensate for the loss of low-SR fibers at CF of ouabain animals and thus encode the TFS via phase-locking even at suprathreshold levels (Carney, 2018). High vector strength, an indicator for phase locking, is not limited to frequencies near CF in the gerbil (Versteegh et al., 2011). Within the framework of a master thesis, single-unit ANF responses to TFS1 stimuli of a subset of the gerbils of this study together with other old and ouabain treated gerbils were investigated (Fandrich, 2020). The results demonstrated, at the stage of ANFs, all gerbil groups can discriminate harmonic from inharmonic tone complexes. However, despite significant differences between the gerbil groups, it could not be excluded that these differences arise from an unequal distribution of ANFs with low characteristic frequencies between groups in which phase-locking and thus temporal coding is strongest. Thus, the observed age-related difference in behavioral TFS1 discrimination cannot be explained by the results of the analysis of single-unit ANF responses. Furthermore, Oetjen et al. (2020) reported no correlation between the number of functional synapses at locations corresponding to 2 and 4 kHz and TFS1 test results for the gerbils of the present study. Thus, central processing deficits, instead of peripheral deficits, could have led to the observed age-related temporal perception deficits. Recent results in humans also showed no association between impaired speech in noise perception and ABR or envelope-following responses, which have been considered as measures for synaptopathy (Guest et al., 2018).

#### **4.5.3 Effects of central processing deficits on temporal perception**

Kale et al. (2014) quantified the representation of TFS cues in ANF responses of normal-hearing chinchillas presented with TFS1 stimuli by analyzing the cross

correlation of neural responses to harmonic and inharmonic tone complexes. They observed, in contrast to the human listener's performance, no dependence on harmonic rank even for unresolved harmonics ( $N=20$ ). Assuming chinchillas have similar TFS1 perception to humans, Kale et al. (2014) concluded that other factors than peripheral processing must have limited the chinchilla's performance. Compromised temporal perception in old gerbils has been suggested to result from a loss of central inhibition (Gleich et al., 2003; Khouri et al. 2011). Kessler et al. (2020) observed an age-related decrease of binding of Flumazenil to GABA receptors in the IC, medial geniculate body and auditory cortex of gerbils. An improvement of impaired temporal processing was observed in old gerbils treated with a blocker of the GABA-degrading enzyme (Gleich et al., 2003). All these findings suggest a possible role of central inhibition for TFS processing.

## **5 General Discussion**

Synaptopathy in humans caused by aging has been demonstrated in a recent study (Wu et al., 2019). Noise-induced synaptopathy in humans and their consequences for perception was studied mainly with electrophysiological and behavioral measures with conflicting results (e.g. Moore et al., 2014; Prendergast et al., 2017a and b; Bramhall et al., 2019). The diagnostic of synaptopathy in humans would be the first step for effective treatment. As a first step, suggested by Plack et al. (2016), it is necessary to rely on animal data in the same species relating comparable electrophysiological and behavioral measures with direct histological measures. This chapter compares the temporal perception in the gerbil outlined in chapters 2 to 4 with a focus on the effect of aging and experimentally induced synaptopathy to electrophysiological and histological measures in the same species. By comparing the results of this thesis to the current state of research on perception of humans, the basis for possible future treatment of compromised temporal perception should be laid. In the following, the findings of each chapter are briefly summarized and joint conclusions are drawn. Finally, important limitations of the results are identified and suggestions for future research are given.

## **5.1 Summary of the experimental chapters**

The goal of chapter 2 was to investigate the relation of temporal perception and its neural representation in young normal-hearing gerbils. Therefore, behavioral discrimination of SCHR complexes was measured and compared to neuronal single-unit responses of SCHR complexes in the AN and IC using four different  $f_0$ s. This study provided the first comparison of behavioral and neural responses to SCHR complexes within the same species. To distinguish between the effects of TFS and ENV on temporal perception, two different experiments were conducted with either only TFS cues being available for discrimination or both acoustical TFS and ENV cues being available. Additionally, the experiment presenting only acoustical TFS cues was conducted in young normal-hearing human participants. The main findings can be summarized as follows: Gerbils and humans were able to discriminate SCHR complexes up to  $f_0 = 200$  Hz in a behavioral paradigm. In the case of only TFS cues being available for discrimination, the correlations between behavioral and neural TFS- and ENV-based sensitivities were high for both AN and IC responses. With TFS and ENV cues available for discrimination, the correlations between behavioral were

highest for either rate- or ENV-based AN sensitivities, and for TFS- and ENV-based IC sensitivities.

The study presented in chapter 3, investigated whether the perception of temporal properties in the Mongolian gerbil is affected by age. Since the measurement, outlined in chapter 2 already provided data of SCHR complex discrimination from young normal-hearing animals, the same stimulus parameters were used for old gerbils. The hearing status of the old gerbils was verified using ABR measurements to click stimuli prior to the behavioral testing, which also served as a predictor for peripheral HL. The results demonstrated that old gerbils can discriminate SCHR complexes, similar to young gerbils, up to  $f_0 = 200$  Hz, but overall achieved lower sensitivities than younger animals. Therefore, discrimination of SCHR complexes was appropriate to detect age-related diminished sensitivity in temporal perception. This finding is in line with age-effects found in the gerbil for other temporal-related tasks, like gap detection (Hamann et al., 2004) or forward masking (Gleich et al., 2007). No differences in sensitivity were observed between young and old gerbils for  $f_0$  below 100 Hz indicating no threshold shift in temporal perception due to aging. The ABRs revealed age-related reduced wave I amplitudes, indicating synaptopathy. Additionally, the observed decrease in behavioral sensitivity with increasing  $f_0$  suggests age-related deficits in central coding. As previously demonstrated, an altered inhibitory network may be responsible for the decreased sensitivity in temporal perception in old animals.

The influence of experimentally induced synaptopathy and aging on temporal perception was investigated in chapter 4. In this experiment, young gerbils were treated with ouabain, with the aim of degenerating the low-SR fibers (Bourien et al., 2014). Young untreated, old and young ouabain treated gerbils discriminated behaviorally harmonic and frequency-shifted inharmonic tone complexes in the TFS1 test (Moore and Sek, 2009). The frequency shift needed for the discrimination threshold of  $d' = 1$  and absolute sensitivities were evaluated. Additionally, measuring ABRs to click stimuli provided an estimate of peripheral HL, like in chapter 3. As shown in the results, ARHL negatively affects sensitivity in temporal perception. No differences were observed within the groups of young gerbils including the ouabain treated groups, suggesting low-SR fiber loss did not affect temporal perception. In general, this observation was made for both paradigms, investigating sensitivities for all frequency shifts and taking only discrimination thresholds into account. The ABR

amplitudes were decreased for all groups with respect to the young untreated gerbils, indicating synaptopathy due to aging and ouabain treatment. The synapse numbers were not reduced for the low ouabain and the sham surgery group whereas the high ouabain group showed reduced numbers of synapses similar to that in old gerbils (Steenken and Bovee, 2020, personal communication). Oetjen et al. (2020) reported no correlation between the number of remaining synapses per IHC and temporal perception in gerbils. Overall, the results indicate, the exclusive loss of low-SR fibers is not sufficient for explaining temporal perception deficits.

## **5.2 Common findings from the three experiments**

Discrimination of SCHR complexes and the TFS1 test indicated good TFS perception of young gerbils. Investigating effects of ARHL on temporal perception in chapters 3 and 4, SCHR complex discrimination and the TFS1 test results revealed decreased sensitivity in temporal perception with increasing age. The ABR wave I amplitudes to click stimuli of old gerbils were also decreased in respect to young gerbils. These findings are consistent and indicate that old gerbils suffer from synaptopathy, which has been demonstrated in previous studies (e.g. Schmiedt et al., 1996; Gleich et al., 2016), and show compromised temporal perception. Whether synaptopathy is the only cause of age-related decreased sensitivity in temporal perception is not proven by the present data since old gerbils suffer from an additional variety of age-related alterations, like stria vascularis dysfunction (Heeringa and Köppl, 2019), central processing deficits (Khouri et al., 2011) or loss of inhibition (Kessler et al., 2020).

In the TFS1 test, ouabain-treated gerbils showed no loss of sensitivity in temporal perception. However, ABRs indicated experimentally induced loss of low-SR fibers. This finding does not contradict the data observed for old gerbils as it only suggests that the loss of low-SR fibers may not be sufficient to decrease TFS perception in young gerbils. Transferring the results of old gerbils in the TFS1 test to the SCHR complex discrimination, points to central deficits being involved in age-related decreased sensitivity. The observed decrease in sensitivity with increasing  $f_0$  in discriminating SCHR complexes also supports this hypothesis since in higher stages of the auditory pathway temporal coding is limited to lower frequencies (Joris et al., 2003)

Differences between SCHR complex discrimination and TFS1 test results were observed when analyzing the threshold changes for temporal perception due to aging.

Young and old gerbils had the same threshold of  $f_0 = 200$  Hz in the SCHR complex discrimination. In contrast to that, in the TFS1 test old gerbils had higher thresholds than young untreated animals. This inconsistency can be explained by the different utilized step sizes of the fundamental frequency (4 steps) and frequency shift (up to 10 steps). Smaller  $f_0$  step sizes in the SCHR complex discrimination in the range slightly above the threshold of young gerbils, between  $f_0 = 100$  and 200 Hz, probably would have revealed an age-related threshold dependency. Other gerbil studies also showed no age-related threshold shift in temporal-related tasks, but observed age effects above threshold (e.g. Hamann et al., 2004). These studies also used big step sizes.

Another difference between the utilized paradigms is the comparison to humans. While young gerbils accomplished lower sensitivities than young normal-hearing humans in the SCHR complex discrimination, the untreated gerbils in the TFS1-test performed similarly to human TFS1 results from the literature (Moore and Sek., 2009; Jackson and Moore, 2014; Marmel et al., 2015). The different procedures used for humans and gerbils may provide an explanation. The human TFS1 test results were conducted in a 2-Down/1-Up 2-alternative forced-choice procedure, with thresholds corresponding to  $d' = 1$ . The human SCHR complex discrimination, however, was conducted in a simulated Go/NoGo procedure to match the procedure in gerbils. The human listeners made, in contrast to the gerbils, close to zero false alarms, with the hit rates similar to the gerbils. The general higher false alarm rates of the gerbils limited their sensitivity. The humans, however, achieved due to their natural lower nervousness higher sensitivities in the SCHR complex discrimination but could not benefit from that advantage in the TFS1 test.

## **5.3 Comparison of temporal perception in gerbils and humans**

### **5.3.1 SCHR complex discrimination in old humans**

A main finding of this thesis was the decreased sensitivity in temporal perception in old gerbils. Similar to that, age-related loss of sensitivity for temporal perception has been observed in humans in a variety of experiments including the TFS1 test (e.g. Plomp, 1989; Pichora-Fuller and Schneider, 1992; Snell, 1997; Strouse et al., 1998; Hopkins and Moore, 2011; Moore et al., 2012). Reduced sensitivity for SCHR complex discrimination was also observed in old hearing-impaired and old cochlea-implant

humans (Lauer et al., 2009; Drennan et al., 2008) with no data of normal-hearing elderly humans being available. Given the strong hints for age-related loss of sensitivity in temporal perception in humans together with the findings of age-related loss of sensitivity in gerbils, it can be concluded that old humans might also show reduced sensitivity in SCHR complex discrimination. A similar suggestion was also made by Lauer et al. (2009).

### **5.3.2 Loss of low-SR fibers with aging in humans**

The age-related reduction in ABR wave I amplitudes to click stimuli suggests that old gerbils suffer from synaptopathy, as was previously shown in the literature (e.g. Schmiedt et al., 1996). Synapse counting in the old gerbils confirmed this suggestion (Steenken and Bovee, 2020, personal communication). The supra-threshold ABR reduction indicated a loss of low-SR fibers (Furman et al., 2013). Compared to medium- and high- SR fibers, these fibers are more vulnerable to aging effects in animals (Schmiedt et al., 1996), noise-exposure (Furman et al., 2013) and ouabain-induced synaptopathy (Bourien et al., 2014). In humans, the loss of OHCs and ANFs caused by aging was revealed in a histological study (Wu et al., 2019). However, no distinction between the fiber types was made since direct measurements of SR would require invasive techniques *in vivo*. A possible indicator for synaptopathy in humans could be reduced ABR wave I to wave V ratio at suprathreshold levels (Sergeyenko et al., 2013; Möhrle et al., 2016; Bharadwaj et al., 2019). Given aging results in similar peripheral changes in gerbils and humans, one could assume from the present data that humans also lose especially low-SR fibers due to aging.

### **5.3.3 Effect of low-SR fibers loss on temporal perception in humans**

It is unclear to what extent synaptopathy (i.e. loss of low-SR fibers) affects temporal perception or perception at all (Bramhall et al., 2019) in humans. In the gerbils, the synapse numbers were not reduced for the low ouabain group whereas the synapse loss of the high ouabain group was similar to that in old gerbils (Steenken and Bovee, 2020, personal communication). However, no reduced temporal perception was observed for the young gerbil groups with ouabain treatment. Oetjen et al. (2020) reported no correlation between the number of remaining synapses per IHC and temporal perception in gerbils. Assuming direct transferability, a loss of low-SR fibers would not affect temporal perception in humans. Similarly, in a model for detecting interaural time differences, based on TFS, a synaptic loss of 50 % was predicted to



affect the sensitivity only by a factor of  $\sqrt{2}$  (Oxenham, 2016). However, this model does not distinguish between the loss of low-SR fibers or medium- and high-SR fibers. Singling out the effects of low-SR fibers on temporal perception is also difficult since the human ANF SR-distribution along the BM is unknown. Peri-stimulus time responses predicted a mean SR of 23 spikes/s in the 4 kHz frequency range in old normal-hearing humans but no measurements were conducted for other frequencies (Huet et al., 2021).

Including not only the specific loss of low-SR fibers but also other features of synaptopathy in humans offers more studies for comparison. Patients with SNHL performed more poorly in the TFS1 test and showed reduced electrophysiological responses (Mathew et al., 2016). No connection was observed between age-related synaptopathy, indicated by electrophysiological measures, and impaired speech in noise perception for listeners with normal audiograms (e.g. Prendergast et al., 2017b; Guest et al., 2018; Johannesen et al., 2019). However, for listeners with elevated thresholds, synaptopathy results in a reduced speech in noise perception (Garret et al., 2020). Overall, the degree of synaptopathy probably determines the reduction in sensitivity in temporal perception in humans. In line with that, the data of the ouabain treated gerbils suggests that an exclusive loss of low-SR fibers is not sufficient to reduce temporal perception in humans.

#### **5.3.4 Effect of altered central inhibition on temporal perception in humans**

The data outlined in chapter 2 lead to the conclusion that temporal features are preserved in the central system up to the IC in young gerbils to some extent. Chapters 3 and 4 indicated that deficits in central processing, altered central inhibition (Gleich et al., 2003; Kessler et al., 2019), could be a major factor for compromised temporal perception in old gerbils. Furthermore, an interplay of peripheral and central deficits was suggested to be responsible for compromised temporal coding, i.e. a loss of ANFs reduces peripheral output, which arises central coding deficits (Heeringa et al., 2020). Similar to gerbils, altered inhibition was revealed with increasing age in humans (Banay-Schwartz et al., 1993; Pal et al., 2019; review: Caspary et al., 2008). Thus, it is likely that also temporal coding of elderly humans is negatively affected by a loss of central inhibition.

## 5.4 Limitations and suggestions for future research

### 5.4.1 Open questions and limitations

In SCHR complex discrimination, the effect of the  $C$  value could not be precisely examined. Young gerbils showed reduced sensitivities for  $C = 0.25$  in the direction experiment, while for old gerbils no effect of  $C$  value was observed. Similarly, in the velocity experiment young gerbils showed reduced sensitivities for small differences in  $C$  between reference and target stimuli. Leek et al. (2004) conducted the velocity experiment in humans and birds using  $C$  values with a step size of  $C = 0.1$  observing a sensitivity reduction mainly for  $C < 0.5$ . To measure more accurately the sensitivity reduction for small  $C$  values,  $C = 0.1$  instead of  $C = 0.75$  could have been also used in the SCHR discrimination with young and old gerbils. Another possible concern of the SCHR direction experiments is a possible loudness cue for discrimination. In humans, the level of SCHR+ complexes has to be up to 5 dB SPL higher than the level of SCHR-complexes to achieve equal loudness for  $f_0 < 400$  Hz (Mauermann and Hohmann, 2007). To exclude a possible loudness cue for discrimination, the stimuli could have been played with a level roving of 10 dB SPL, similar to Leek et al. (2004). However, this procedure would have increased the complexity of the task, and therefore could have led to prolonged training periods in the gerbil. The sensitivity in the TFS1 test, on the other hand, does not depend on the stimulus level in humans (Moore and Sek., 2009), which makes a loudness cue unlikely to play a role in the corresponding gerbil experiment.

The synapse numbers per IHC were counted at different corresponding frequencies for all gerbils who participated in the TFS1 test (Steenken and Bovee, 2020, personal communication). However, synapse counting does not distinguish between the fibers SR. It could not be verified whether the high ouabain group lost exclusively low-SR fibers or whether medium and high SR fibers were additionally degenerated (Bourien et al., 2014). A fiber tracing analysis could evaluate the ANF loss separately for each fiber type, taking advantage of the differences in thickness and side of innervation between the fiber types (Liberman, 1982; review: Heil and Peterson, 2015). First, the fibers have to be characterized electrophysiologically, afterwards stained in vivo and subsequently, their position on the IHC has to be analyzed post mortem.

Additional insights into the effects of synaptopathy on temporal perception could be gained by comparing behavioral sensitivity in the TFS1 test with electrophysiological recordings of the same stimuli in the ANF. Friederike Steenken conducted such recordings with the gerbils used for the behavioral test. The comparison of both data sets would provide additional insight into to what extent TFS information is preserved at the stage of the ANF with or without ouabain treatment.

## **5.4.2 Future Experiments**

### **5.4.2.1 Do humans lose low-SR fibers due to aging?**

The results in chapter 4 suggest that a loss of low-SR fibers may not diminish temporal perception in gerbils. Comparable experiments cannot be conducted in humans since ANF-related research in humans relies on non-invasive techniques in nearly all studies. As shown in gerbils, mass potential recordings at the round window to band-limited noise bursts enable the detection of low-SR fiber loss (Batrel et al., 2017). In this method, a neural gain index is derived, as a ratio between stimulated and unstimulated amplitude spectrum density. For ouabain treated gerbils, the neural gain index decreases for high-frequency noise bursts. Since old humans suffer from ANF loss in contrast to young humans (Wu et al., 2019), investigating the neural index gain in young and old normal-hearing humans, who undergo surgery on the ear anyway, would show a possible age-related loss of low-SR fibers.

### **5.4.2.2 Are high-SR fibers able to compensate for a loss of low-SR fibers?**

The general assumption of medium and high-SR fibers encoding stimuli at the hearing threshold and low-SR fibers encoding at supra-threshold levels is only valid at their CF (Lieberman, 1978). Off-CF, stimuli at increased levels could be encoded by all fiber types. Thus, one proposed explanation for the lack of reduced sensitivity for the ouabain treated gerbils in the TFS1 test could be that off-CF medium and high-SR fibers are sufficient for TFS coding. An AN model for high-SR fibers proposed, the temporal fluctuation in discharge rate differs for vowels at formant and less weighted frequencies (Carney, 2018). The profile of temporal fluctuation would code spectral properties of vowels although the average discharge rates are saturated across all channels. To test this hypothesis in gerbils, AN responses to vowels of high SR fibers have to be measured. By comparing the temporal fluctuation in discharge rate at the format frequencies with the fluctuation at other frequencies, the possible contribution

of high-SR fibers to temporal coding at supra-threshold levels can be examined. This experiment could explain the lack of sensitivity reduction of the ouabain-treated gerbils in the TFS1 test.

#### **5.4.2.3 Are speech stimuli more suitable to detect a loss of sensitivity in temporal perception?**

Possibly, the TFS1 test was not sensitive enough to detect the effect of ouabain treatment on temporal perception. In young humans with synaptopathy, no effect of altered temporal perception was observed in all tested non-speech measures (Gordon-Salant and Fitzgibbons, 1999). On the other hand, speech-related measurements in quiet and in noise in the same listeners showed differences in speech recognition performance depending on hearing status. Thus, possibly with measures based on more complex stimuli, like speech, reduced temporal perception caused by ouabain treatment could be revealed in the gerbil. Jüchter and Klump (2021) showed that young normal-hearing gerbils can discriminate logatomes (consonant-vowel-consonant and vowel-consonant-vowel combination; Meyer et al., 2010) in noise with speech-like spectral properties (Dreschler et al., 2001). By measuring sensitivities based on hit and false alarm rates for the most sensitive subset of logatome in noise discriminations for different signal-to-noise ratios with young untreated and young ouabain treated gerbils, the effect of ouabain treatment on speech in noise discrimination could be observed.

#### **5.4.2.4 Is age-related change of sensitivity in temporal perception affected by altered central inhibition?**

Altered inhibition in the central nervous system during the process of aging has been suggested in chapters 3 and 4 to play a key role in the development of reduced sensitivity in temporal perception possibly also in the absence of peripheral pathologies. In general, temporal processing of complex stimuli like speech involves large participation of the auditory cortex (e.g. Zatorre et al., 2002; Cardin, 2018). Aging has been linked to alterations in the inhibitory circuit of the auditory cortex (Kotak et al., 2005; Peelle and Wingfield, 2016; Rogalla and Hildebrandt, 2020). Thus, a possible connection between age-related reduced temporal perception and altered inhibition can be most likely observed in the auditory cortex. To test this hypothesis, the previously suggested logatome experiment (Jüchter and Klump, 2021) could be conducted with old gerbils as well. By using old gerbils without elevated ABR wave I thresholds and

minimal reduced supra-threshold amplitudes, the possible influence of age-related synaptopathy on speech discrimination could be reduced (Pittmann-Polletta et al., 2021). Next, the inhibition, either parvalbumin immunoreactivity (Rogalla and Hildebrandt, 2020) or investigation of binding of flumazenil to GABA receptors (Kessler et al. 2020), has to be evaluated in the auditory cortex of the old and young untreated animals. By comparing temporal perception of speech stimuli and inhibition in the auditory cortex, a possible cause of an age-related decline in sensitivity in temporal perception could be identified.

## **5.5 Final conclusions**

In summary, this thesis investigated temporal perception in the gerbil and its representation in the auditory pathway with a focus on the effects of aging and synaptopathy. The results highlighted aging results in compromised temporal perception, whereas synaptopathy seems not. The findings suggest that age-related changes in temporal representation in higher stages of the auditory pathway may explain temporal perception deficits. Future research on the treatment of deficits in temporal perception should focus on the central system as altered inhibition is suggested to be a major factor for compromised temporal perception.

## **6 Appendix**

## 6.1 References of Chapters 1 to 6

- A. Meijer, J. Kleinhuis, G. Drost, M. Van Dijk, P. V. D. (2021). Phase-Locking of the Human Cochlear Nerve Assessed by Direct Electrophysiological Recording During Brain Surgery. *The Association for Research in Otolaryngology (ARO) The 44th Annual MidWinter Meeting*.
- Alberti, P. W. (1992). Noise induced hearing loss. *British Medical Journal*, *304*(6826), 522. <https://doi.org/10.1136/bmj.304.6826.522>
- Aminoff, M. J., & Daroff, R. B. (2014). Encyclopedia of the Neurological Sciences. In *Encyclopedia of the Neurological Sciences*. <https://doi.org/10.1097/01.wno.0000176626.56739.d4>
- Anderson, D. J., Rose, J. E., Hind, J. E., & Brugge, J. F. (1971). Temporal Position of Discharges in Single Auditory Nerve Fibers within the Cycle of a Sine-Wave Stimulus: Frequency and Intensity Effects. *The Journal of the Acoustical Society of America*, *49*(4B). <https://doi.org/10.1121/1.1912474>
- Banay-Schwartz, M., Palkovits, M., & Lajtha, A. (1993). Heterogeneous distribution of functionally important amino acids in brain areas of adult and aging humans. *Neurochemical Research*, *18*(4), 417–423. <https://doi.org/10.1007/BF00967245>
- Barlow, H. B., Levick, W. R., & Yoon, M. (1971). Responses to single quanta of light in retinal ganglion cells of the cat. *Vision Research*, *11*(SUPPL. 3). [https://doi.org/10.1016/0042-6989\(71\)90033-2](https://doi.org/10.1016/0042-6989(71)90033-2)
- Barsz, K., Ison, J. R., Snell, K. B., & Walton, J. P. (2002). Behavioral and neural measures of auditory temporal acuity in aging humans and mice. *Neurobiology of Aging*, *23*(4). [https://doi.org/10.1016/S0197-4580\(02\)00008-8](https://doi.org/10.1016/S0197-4580(02)00008-8)
- Batrel, C., Huet, A., Hasselmann, F., Wang, J., Desmadryl, G., Nouvian, R., Puel, J. L., & Bourien, J. (2017). Mass potentials recorded at the round window enable the detection of low spontaneous rate fibers in gerbil auditory nerve. *PLoS ONE*, *12*(1), 1–16. <https://doi.org/10.1371/journal.pone.0169890>
- Beutelmann, R., Laumen, G., Tollin, D., & Klump, G. M. (2015). Amplitude and phase equalization of stimuli for click evoked auditory brainstem responses. *The Journal of the Acoustical Society of America*, *137*(1), EL71–EL77. <https://doi.org/10.1121/1.4903921>
- Bharadwaj, H. M., Mai, A. R., Simpson, J. M., Choi, I., Heinz, M. G., & Shinn-Cunningham, B. G. (2019). Non-Invasive Assays of Cochlear Synaptopathy – Candidates and Considerations. *Neuroscience*, *407*. <https://doi.org/10.1016/j.neuroscience.2019.02.031>
- Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *Journal of Neuroscience*, *35*(5), 2161–2172. <https://doi.org/10.1523/JNEUROSCI.3915-14.2015>
- Bharadwaj, H. M., Verhulst, S., Shaheen, L., Charles Liberman, M., & Shinn-

- Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound. *Frontiers in Systems Neuroscience*, 8(FEB), 1–18.  
<https://doi.org/10.3389/fnsys.2014.00026>
- Boettcher, F. A., Mills, J. H., Swerdloff, J. L., & Holley, B. L. (1996). Auditory evoked potentials in aged gerbils: Responses elicited by noises separated by a silent gap. *Hearing Research*, 102(1–2), 167–178. [https://doi.org/10.1016/S0378-5955\(96\)90016-7](https://doi.org/10.1016/S0378-5955(96)90016-7)
- Boettcher, Flint A. (2002). Susceptibility to acoustic trauma in young and aged gerbils. *The Journal of the Acoustical Society of America*, 112(6), 2948–2955.  
<https://doi.org/10.1121/1.1513364>
- Bone, R. C. (1978). Noise-induced threshold shift and cochlear pathology in the mongolian gerbil. *Journal of the Acoustical Society of America*, 63(4).  
<https://doi.org/10.1121/1.381822>
- Bourien, J., Tang, Y., Batrel, C., Huet, A., Lenoir, M., Ladrech, S., Desmadryl, G., Nouvian, R., Puel, J. L., & Wang, J. (2014). Contribution of auditory nerve fibers to compound action potential of the auditory nerve. *Journal of Neurophysiology*, 112(5), 1025–1039. <https://doi.org/10.1152/jn.00738.2013>
- Bramhall, N., Beach, E. F., Epp, B., Le Prell, C. G., Lopez-Poveda, E. A., Plack, C. J., Schaette, R., Verhulst, S., & Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing Research*, 377, 88–103. <https://doi.org/10.1016/j.heares.2019.02.016>
- Bramhall, N. F., Konrad-Martin, D., McMillan, G. P., & Griest, S. E. (2017). Auditory Brainstem Response Altered in Humans with Noise Exposure Despite Normal Outer Hair Cell Function. *Ear and Hearing*, 38(1).  
<https://doi.org/10.1097/AUD.0000000000000370>
- Bredberg, G. (1968). Cochlear structure and hearing in man. *Ciba Foundation Symposium-Hearing Mechanisms in Vertebrates*, 126–142.
- Brownell, W. E. (1983). *Observation on a motile response in isolated outer hair cells. Mechanisms of Hearing*. Monash University Press, Clayton, Australia.
- Brugge, J. F., Volkov, I. O., Garell, P. C., Reale, R. A., & Howard, M. A. (2003). Functional Connections between Auditory Cortex on Heschl's Gyrus and on the Lateral Superior Temporal Gyrus in Humans. *Journal of Neurophysiology*, 90(6).  
<https://doi.org/10.1152/jn.00500.2003>
- Brughera, A., Dunai, L., & Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: The high-frequency limit. *The Journal of the Acoustical Society of America*, 133(5). <https://doi.org/10.1121/1.4795778>
- Burns, J. C., & Corwin, J. T. (2013). A historical to present-day account of efforts to answer the question: “What puts the brakes on mammalian hair cell regeneration?” In *Hearing Research* (Vol. 297).  
<https://doi.org/10.1016/j.heares.2013.01.005>



- Buus, S., Klump, G. M., Gleich, O., & Langemann, U. (1995). An excitation-pattern model for the starling (*Sturnus vulgaris*). *Journal of the Acoustical Society of America*, 98(1), 112–124. <https://doi.org/10.1121/1.414466>
- Carcagno, S., & Plack, C. J. (2020). Effects of age on electrophysiological measures of cochlear synaptopathy in humans. *Hearing Research*, 396, 108068. <https://doi.org/10.1016/j.heares.2020.108068>
- Carcagno, S., & Plack, C. J. (2021). Effects of age on psychophysical measures of auditory temporal processing and speech reception at low and high levels. *Hearing Research*, 400, 108117. <https://doi.org/10.1016/j.heares.2020.108117>
- Cardin, J. A. (2018). Inhibitory Interneurons Regulate Temporal Precision and Correlations in Cortical Circuits. In *Trends in Neurosciences* (Vol. 41, Issue 10). <https://doi.org/10.1016/j.tins.2018.07.015>
- Cariani, P. (1999). Temporal coding of periodicity pitch in the auditory system: An overview. *Neural Plasticity*, 6(4). <https://doi.org/10.1155/NP.1999.147>
- Carlyon, R. P., & Datta, A. J. (1997). Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system. *The Journal of the Acoustical Society of America*, 101(6), 3636–3647. <https://doi.org/10.1121/1.418324>
- Carlyon, R. P., Flanagan, S., & Deeks, J. M. (2017). A Re-examination of the Effect of Masker Phase Curvature on Non-simultaneous Masking. *JARO - Journal of the Association for Research in Otolaryngology*, 18(6), 815–825. <https://doi.org/10.1007/s10162-017-0637-5>
- Carney, L. H. (2018). Supra-Threshold Hearing and Fluctuation Profiles: Implications for Sensorineural and Hidden Hearing Loss. *JARO - Journal of the Association for Research in Otolaryngology*, 19(4), 331–352. <https://doi.org/10.1007/s10162-018-0669-5>
- Carney, L. H., McDuffy, M. J., & Shekhter, I. (1999). Frequency glides in the impulse responses of auditory-nerve fibers. *The Journal of the Acoustical Society of America*, 105(4). <https://doi.org/10.1121/1.426843>
- Carney, L. H., Zilany, M. S. A., Huang, N. J., Abrams, K. S., & Idrobo, F. (2014). Suboptimal use of neural information in a mammalian auditory system. *Journal of Neuroscience*, 34(4). <https://doi.org/10.1523/JNEUROSCI.3031-13.2014>
- Caspary, D. M., Raza, A., Lawhorn Armour, B. A., Pippin, J., & Arneric, S. P. (1990). Immunocytochemical and neurochemical evidence for age-related loss of GABA in the inferior colliculus: Implications for neural presbycusis. *Journal of Neuroscience*, 10(7), 2363–2372. <https://doi.org/10.1523/jneurosci.10-07-02363.1990>
- Caspary, Donald M., Ling, L., Turner, J. G., & Hughes, L. F. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *Journal of Experimental Biology*, 211(11), 1781–1791.

<https://doi.org/10.1242/jeb.013581>

- Caspary, Donald M., & Llano, D. A. (2019). Aging Processes in the Subcortical Auditory System. In *The Oxford Handbook of the Auditory Brainstem*.  
<https://doi.org/10.1093/oxfordhb/9780190849061.013.16>
- Caspary, Donald M., Milbrandt, J. C., & Helfert, R. H. (1995). Central auditory aging: GABA changes in the inferior colliculus. *Experimental Gerontology*, 30(3–4), 349–360. [https://doi.org/10.1016/0531-5565\(94\)00052-5](https://doi.org/10.1016/0531-5565(94)00052-5)
- Cedolin, L., & Delgutte, B. (2010). Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve. *Journal of Neuroscience*, 30(38), 12712–12724. <https://doi.org/10.1523/JNEUROSCI.6365-09.2010>
- Cheal, M. Lou. (1986). The gerbil: A unique model for research on aging. *Experimental Aging Research*, 12(1), 3–21. <https://doi.org/10.1080/03610738608259430>
- Costalupes, J. A., Young, E. D., & Gibson, D. J. (1984). Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat. *Journal of Neurophysiology*, 51(6), 1326–1344. <https://doi.org/10.1152/jn.1984.51.6.1326>
- Dau, T., Wegner, O., Mellert, V., & Kollmeier, B. (2000). Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *The Journal of the Acoustical Society of America*, 107(3), 1530–1540.  
<https://doi.org/10.1121/1.428438>
- Davis, K. A., Ding, J., Benson, T. E., & Voigt, H. F. (1996). Response properties of units in the dorsal cochlear nucleus of unanesthetized decerebrate gerbil. *Journal of Neurophysiology*, 75(4). <https://doi.org/10.1152/jn.1996.75.4.1411>
- De Boer, E. (1956). Pitch of inharmonic signals [4]. In *Nature* (Vol. 178, Issue 4532).  
<https://doi.org/10.1038/178535a0>
- de Cheveigné, A., & Pressnitzer, D. (2006). The case of the missing delay lines: Synthetic delays obtained by cross-channel phase interaction. *The Journal of the Acoustical Society of America*, 119(6). <https://doi.org/10.1121/1.2195291>
- Devore, S., & Delgutte, B. (2010). Effects of reverberation on the directional sensitivity of auditory neurons across the tonotopic axis: Influences of interaural time and level differences. *Journal of Neuroscience*, 30(23), 7826–7837.  
<https://doi.org/10.1523/JNEUROSCI.5517-09.2010>
- Divenyi, P. L., Stark, P. B., & Haupt, K. M. (2005). Decline of speech understanding and auditory thresholds in the elderly. *The Journal of the Acoustical Society of America*, 118(2), 1089–1100. <https://doi.org/10.1121/1.1953207>
- Dolphin, W. F., & Mountain, D. C. (1992). The envelope following response: Scalp potentials elicited in the mongolian gerbil using sinusoidally AM acoustic signals. *Hearing Research*, 58(1), 70–78. [https://doi.org/10.1016/0378-5955\(92\)90010-K](https://doi.org/10.1016/0378-5955(92)90010-K)
- Dooling, R. J., Leek, M. R., Gleich, O., & Dent, M. L. (2002). Auditory temporal resolution in birds: Discrimination of harmonic complexes. *The Journal of the Acoustical Society of America*, 112(2), 748–759.

<https://doi.org/10.1121/1.1494447>

- Drennan, W. R., Longnion, J. K., Ruffin, C., & Rubinstein, J. T. (2008). Discrimination of schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. *JARO - Journal of the Association for Research in Otolaryngology*, 9(1), 138–149. <https://doi.org/10.1007/s10162-007-0107-6>
- Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. *International Journal of Audiology*, 40(3). <https://doi.org/10.3109/00206090109073110>
- Dreyer, A., & Delgutte, B. (2006). Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization. *Journal of Neurophysiology*, 96(5). <https://doi.org/10.1152/jn.00326.2006>
- Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *Journal of the Acoustical Society of America*, 97(1). <https://doi.org/10.1121/1.413112>
- Eipert, L., & Klump, G. M. (2020). Uncertainty-based informational masking in a vowel discrimination task for young and old Mongolian gerbils. *Hearing Research*, 392. <https://doi.org/10.1016/j.heares.2020.107959>
- Eipert, L., Selle, A., & Klump, G. M. (2019). Uncertainty in location, level and fundamental frequency results in informational masking in a vowel discrimination task for young and elderly subjects. *Hearing Research*, 377, 142–152. <https://doi.org/10.1016/j.heares.2019.03.015>
- Escanilla, O. D., Victor, J. D., & Di Lorenzo, P. M. (2015). Odor-taste convergence in the nucleus of the solitary tract of the awake freely licking rat. *Journal of Neuroscience*, 35(16). <https://doi.org/10.1523/JNEUROSCI.3526-14.2015>
- Evans, E. F. (1975). The sharpening of cochlear frequency selectivity in the normal and abnormal cochlea. *International Journal of Audiology*, 14(5–6). <https://doi.org/10.3109/00206097509071754>
- Evans, E. F., Pratt, S. R., Spenner, H., & Cooper, N. P. (1992). Comparisons of Physiological and Behavioural Properties: Auditory Frequency Selectivity. In *Auditory Physiology and Perception*. <https://doi.org/10.1016/b978-0-08-041847-6.50024-1>
- Faulstich, M., & Kössl, M. (1999). Neuronal response to cochlear distortion products in the anteroventral cochlear nucleus of the gerbil. *The Journal of the Acoustical Society of America*, 105(1), 491–502. <https://doi.org/10.1121/1.424586>
- Fischer, N., Johnson Chacko, L., Glueckert, R., & Schrott-Fischer, A. (2020). Age-Dependent Changes in the Cochlea. In *Gerontology* (Vol. 66, Issue 1). <https://doi.org/10.1159/000499582>
- Frisina, S. T., Mapes, F., Kim, S. H., Frisina, D. R., & Frisina, R. D. (2006). Characterization of hearing loss in aged type II diabetics. *Hearing Research*,

211(1–2). <https://doi.org/10.1016/j.heares.2005.09.002>

- Fujioka, M., Okano, H., & Edge, A. S. B. (2015). Manipulating cell fate in the cochlea: A feasible therapy for hearing loss. In *Trends in Neurosciences* (Vol. 38, Issue 3). <https://doi.org/10.1016/j.tins.2014.12.004>
- Fulbright, A. N. C., Le Prell, C. G., Griffiths, S. K., & Lobarinas, E. (2017). Effects of Recreational Noise on Threshold and Suprathreshold Measures of Auditory Function. *Seminars in Hearing*, 38(4). <https://doi.org/10.1055/s-0037-1606325>
- Füllgrabe, C., Harland, A. J., Şek, A. P., & Moore, B. C. J. (2017). Development of a method for determining binaural sensitivity to temporal fine structure. *International Journal of Audiology*, 56(12), 926–935. <https://doi.org/10.1080/14992027.2017.1366078>
- Füllgrabe, C., & Moore, B. C. J. (2018). The Association Between the Processing of Binaural Temporal-Fine-Structure Information and Audiometric Threshold and Age: A Meta-Analysis. *Trends in Hearing*, 22, 1–14. <https://doi.org/10.1177/2331216518797259>
- Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, 7(JAN), 1–25. <https://doi.org/10.3389/fnagi.2014.00347>
- Füllgrabe, C., Şek, A. P., & Moore, B. C. J. (2018). Senescent Changes in Sensitivity to Binaural Temporal Fine Structure. *Trends in Hearing*, 22, 1–16. <https://doi.org/10.1177/2331216518788224>
- Furman, A. C., Kujawa, S. G., & Charles Liberman, M. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of Neurophysiology*, 110(3), 577–586. <https://doi.org/10.1152/jn.00164.2013>
- Gallun, F. J., McMillan, G. P., Molis, M. R., Kampel, S. D., Dann, S. M., & Konrad-Martin, D. L. (2014). Relating age and hearing loss to monaural, bilateral, and binaural temporal sensitivity. *Frontiers in Neuroscience*, 8(8 JUN), 1–14. <https://doi.org/10.3389/fnins.2014.00172>
- Garrett, M., Vasilkov, V., Mauermann, M., Wilson, J. L., Henry, K. S., & Verhulst, S. (2020). Speech-in-noise intelligibility difficulties with age: the role of cochlear synaptopathy. *BioRxiv*.
- Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 47(1–2). [https://doi.org/10.1016/0378-5955\(90\)90170-T](https://doi.org/10.1016/0378-5955(90)90170-T)
- Gleich, O., Hamann, I., Kittel, M. C., Klump, G. M., & Strutz, J. (2007). Forward masking in gerbils: The effect of age. *Hearing Research*, 223(1–2), 122–128. <https://doi.org/10.1016/j.heares.2006.11.001>
- Gleich, O., Hamann, I., Klump, G. M., Kittel, M., & Strutz, J. (2003). Boosting GABA improves impaired auditory temporal resolution in the gerbil. *NeuroReport*, 14(14),

1877–1880. <https://doi.org/10.1097/00001756-200310060-00024>

- Gleich, O., Kittel, M. C., Klump, G. M., & Strutz, J. (2007). Temporal integration in the gerbil: The effects of age, hearing loss and temporally unmodulated and modulated speech-like masker noises. *Hearing Research*, *224*(1–2), 101–114. <https://doi.org/10.1016/j.heares.2006.12.002>
- Gleich, O., Leek, M., & Dooling, R. (2007). Influence of Neural Synchrony on the Compound Action Potential, Masking, and the Discrimination of Harmonic Complexes in Several Avian and Mammalian Species. In *Hearing – From Sensory Processing to Perception*. [https://doi.org/10.1007/978-3-540-73009-5\\_1](https://doi.org/10.1007/978-3-540-73009-5_1)
- Gleich, O., Semmler, P., & Strutz, J. (2016). Behavioral auditory thresholds and loss of ribbon synapses at inner hair cells in aged gerbils. *Experimental Gerontology*, *84*, 61–70. <https://doi.org/10.1016/j.exger.2016.08.011>
- Gleich, O., & Strutz, J. (2011). The effect of gabapentin on gap detection and forward masking in young and old gerbils. *Ear and Hearing*, *32*(6), 741–749. <https://doi.org/10.1097/AUD.0b013e318222289f>
- Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech and Hearing Research*, *36*(6), 1276–1285. <https://doi.org/10.1044/jshr.3606.1276>
- Gordon-Salant, Sandra, & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech, Language, and Hearing Research*, *42*(2), 300–311. <https://doi.org/10.1044/jslhr.4202.300>
- Gratton, M. A., & Schulte, B. A. (1995). Alterations in microvasculature are associated with atrophy of the stria vascularis in quiet-aged gerbils. *Hearing Research*, *82*(1), 44–52. [https://doi.org/10.1016/0378-5955\(94\)00161-I](https://doi.org/10.1016/0378-5955(94)00161-I)
- Green, D. M., Swets, J. A., & others. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.
- Greenwood, D. D. (1961). Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane. *The Journal of the Acoustical Society of America*, *33*(10), 1344–1356. <https://doi.org/10.1121/1.1908437>
- Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *Journal of the Acoustical Society of America*, *87*(6), 2592–2605. <https://doi.org/10.1121/1.399052>
- Grose, J. H., Buss, E., & Hall, J. W. (2017). Loud Music Exposure and Cochlear Synaptopathy in Young Adults: Isolated Auditory Brainstem Response Effects but No Perceptual Consequences. *Trends in Hearing*, *21*. <https://doi.org/10.1177/2331216517737417>
- Grose, J. H., & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear and Hearing*, *31*(6), 755–760. <https://doi.org/10.1097/AUD.0b013e3181e627e7>
- Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018).

- Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing Research*, 364, 142–151. <https://doi.org/10.1016/j.heares.2018.03.008>
- Hamann, I., Gleich, O., Klump, G. M., Kittel, M. C., & Strutz, J. (2004). Age-Dependent Changes of Gap Detection in the Mongolian Gerbil (*Meriones unguiculatus*). *JARO - Journal of the Association for Research in Otolaryngology*, 5(1), 49–57. <https://doi.org/10.1007/s10162-003-3041-2>
- Harrison, R. V., & Evans, E. F. (1979). Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig. *Archives of Oto-Rhino-Laryngology*, 224(1–2). <https://doi.org/10.1007/BF00455226>
- Heeringa, A. N., & Köppl, C. (2019). The aging cochlea: Towards unraveling the functional contributions of strial dysfunction and synaptopathy. *Hearing Research*, 376, 111–124. <https://doi.org/10.1016/j.heares.2019.02.015>
- Heeringa, A. N., Zhang, L., Ashida, G., Beutelmann, R., Steenken, F., & Köppl, C. (2020). Temporal coding of single auditory nerve fibers is not degraded in aging gerbils. *Journal of Neuroscience*, 40(2), 343–354. <https://doi.org/10.1523/JNEUROSCI.2784-18.2019>
- Heil, P., Neubauer, H., Irvine, D. R. F., & Brown, M. (2007). Spontaneous activity of auditory-nerve fibers: Insights into stochastic processes at ribbon synapses. *Journal of Neuroscience*, 27(31). <https://doi.org/10.1523/JNEUROSCI.1512-07.2007>
- Heil, P., & Peterson, A. J. (2015). Basic response properties of auditory nerve fibers: a review. *Cell and Tissue Research*, 361(1), 129–158. <https://doi.org/10.1007/s00441-015-2177-9>
- Heinz, M. G., & Young, E. D. (2004). Response Growth with Sound Level in Auditory-Nerve Fibers after Noise-Induced Hearing Loss. *Journal of Neurophysiology*, 91(2). <https://doi.org/10.1152/jn.00776.2003>
- Henderson, D., Bielefeld, E. C., Harris, K. C., & Hu, B. H. (2006). The role of oxidative stress in noise-induced hearing loss. *Ear and Hearing*, 27(1), 1–19. <https://doi.org/10.1097/01.aud.0000191942.36672.f3>
- Henry, K. S., & Heinz, M. G. (2012). Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nature Neuroscience*, 15(10), 1362–1364. <https://doi.org/10.1038/nn.3216>
- Henry, K. S., & Heinz, M. G. (2013). Effects of sensorineural hearing loss on temporal coding of narrowband and broadband signals in the auditory periphery. In *Hearing Research* (Vol. 303). <https://doi.org/10.1016/j.heares.2013.01.014>
- Henry, K. S., Kale, S., & Heinz, M. G. (2016). Distorted tonotopic coding of temporal envelope and fine structure with noise-induced hearing loss. *Journal of Neuroscience*, 36(7), 2227–2237. <https://doi.org/10.1523/JNEUROSCI.3944-16.2016>

15.2016

- Henry, X. K. S., Sayles, M., Hickox, X. A. E., & Heinz, X. M. G. (2019). *Common Etiologies of Sensorineural Hearing Loss*. 39(35), 6879–6887.
- Hind, S. E., Haines-Bazrafshan, R., Benton, C. L., Brassington, W., Towle, B., & Moore, D. R. (2011). Prevalence of clinical referrals having hearing thresholds within normal limits. *International Journal of Audiology*, 50(10), 708–716. <https://doi.org/10.3109/14992027.2011.582049>
- Hopkins, K., & Moore, B. C. J. (2010). The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, 127(3), 1595–1608. <https://doi.org/10.1121/1.3293003>
- Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, 130(1), 334–349. <https://doi.org/10.1121/1.3585848>
- Huet, A., Batrel, C., Wang, J., Desmadryl, G., Nouvian, R., Puel, J. L., & Bourien, J. (2019). Sound Coding in the Auditory Nerve: From Single Fiber Activity to Cochlear Mass Potentials in Gerbils. *Neuroscience*, 407, 83–92. <https://doi.org/10.1016/j.neuroscience.2018.10.010>
- Huet, Antoine, Batrel, C., Tang, Y., Desmadryl, G., Wang, J., Puel, J. L., & Bourien, J. (2016). Sound coding in the auditory nerve of gerbils. *Hearing Research*, 338, 32–39. <https://doi.org/10.1016/j.heares.2016.05.006>
- ISO: 389-7:2019, I. (2019). Acoustics. *Reference Zero for the Calibration of Audiometric Equipment — Part 7: Reference Threshold of Hearing under Free-Field and Diffuse-Field Listening Conditions*. *Stand Int Organ:11*.
- Jackson, H. M., & Moore, B. C. J. (2014). The role of excitation-pattern, temporal-fine-structure, and envelope cues in the discrimination of complex tones. *The Journal of the Acoustical Society of America*, 135(3), 1356–1370. <https://doi.org/10.1121/1.4864306>
- Johannesen, P. T., Buzo, B. C., & Lopez-Poveda, E. A. (2019). Evidence for age-related cochlear synaptopathy in humans unconnected to speech-in-noise intelligibility deficits. *Hearing Research*, 374, 35–48. <https://doi.org/10.1016/j.heares.2019.01.017>
- John, A. B., Hall, J. W., & Kreisman, B. M. (2012). Effects of advancing age and hearing loss on Gaps-in-Noise test performance. *American Journal of Audiology*, 21(2). [https://doi.org/10.1044/1059-0889\(2012/11-0023\)](https://doi.org/10.1044/1059-0889(2012/11-0023))
- Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *Journal of the Acoustical Society of America*, 68(4). <https://doi.org/10.1121/1.384982>
- Johnson, J. S., Yin, P., O'Connor, K. N., & Sutter, M. L. (2012). Ability of primary

- auditory cortical neurons to detect amplitude modulation with rate and temporal codes: Neurometric analysis. *Journal of Neurophysiology*, 107(12).  
<https://doi.org/10.1152/jn.00812.2011>
- Joris, P. X., Carney, L. H., Smith, P. H., & Yin, T. C. T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *Journal of Neurophysiology*, 71(3).  
<https://doi.org/10.1152/jn.1994.71.3.1022>
- Joris, P. X., Smith, P. H., & Yin, T. C. T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *Journal of Neurophysiology*, 71(3).  
<https://doi.org/10.1152/jn.1994.71.3.1037>
- Joris, P X, Schreiner, C. E., & Rees, A. (2004). *Neural Processing of Amplitude-Modulated Sounds*. 541–577.
- Joris, Philip X. (2003). Interaural time sensitivity dominated by cochlea-induced envelope patterns. *Journal of Neuroscience*, 23(15), 6345–6350.  
<https://doi.org/10.1523/jneurosci.23-15-06345.2003>
- Joris, Philip X., & Yin, T. C. T. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *Journal of the Acoustical Society of America*, 91(1), 215–232. <https://doi.org/10.1121/1.402757>
- Jüchter, C., & Klump, G. M. (2021). Speech Sound Discrimination by Mongolian Gerbils. *The Association for Research in Otolaryngology (ARO) The 44th Annual MidWinter Meeting*.
- Kale, S., & Heinz, M. G. (2010). Envelope coding in auditory nerve fibers following noise-induced hearing loss. *JARO - Journal of the Association for Research in Otolaryngology*, 11(4). <https://doi.org/10.1007/s10162-010-0223-6>
- Kale, S., & Heinz, M. G. (2012). Temporal modulation transfer functions measured from auditory-nerve responses following sensorineural hearing loss. *Hearing Research*, 286(1–2). <https://doi.org/10.1016/j.heares.2012.02.004>
- Kale, S., Micheyl, C., & Heinz, M. G. (2014). Implications of within-fiber temporal coding for perceptual studies of F0 discrimination and discrimination of harmonic and inharmonic tone complexes. *JARO - Journal of the Association for Research in Otolaryngology*, 15(3), 465–482. <https://doi.org/10.1007/s10162-014-0451-2>
- Keine, C., & Rübsamen, R. (2015). Inhibition shapes acoustic responsiveness in spherical bushy cells. *Journal of Neuroscience*, 35(22).  
<https://doi.org/10.1523/JNEUROSCI.0133-15.2015>
- Kessler, M., Mamach, M., Beutelmann, R., Lukacevic, M., Eilert, S., Bascuñana, P., Fasel, A., Bengel, F. M., Bankstahl, J. P., Ross, T. L., Klump, G. M., & Berding, G. (2020). GABAA Receptors in the Mongolian Gerbil: a PET Study Using [18F]Flumazenil to Determine Receptor Binding in Young and Old Animals. *Molecular Imaging and Biology*, 22(2), 335–347. <https://doi.org/10.1007/s11307->



- Khanna, S. M., & Leonard, D. G. B. (1982). Basilar membrane tuning in the cat cochlea. *Science*, *215*(4530). <https://doi.org/10.1126/science.7053580>
- Khouri, L., Lesica, N. A., & Grothe, B. (2011). Impaired auditory temporal selectivity in the inferior colliculus of aged mongolian gerbils. *Journal of Neuroscience*, *31*(27), 9958–9970. <https://doi.org/10.1523/JNEUROSCI.4509-10.2011>
- Kim, D. O., Carney, L., & Kuwada, S. (2020). Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. *Journal of Neurophysiology*, *124*(4). <https://doi.org/10.1152/jn.00279.2020>
- King, A., Hopkins, K., & Plack, C. J. (2014). The effects of age and hearing loss on interaural phase difference discrimination. *The Journal of the Acoustical Society of America*, *135*(1), 342–351. <https://doi.org/10.1121/1.4838995>
- Kittel, M., Wagner, E., & Klump, G. M. (2002). An estimate of the auditory-filter bandwidth in the Mongolian gerbil. *Hearing Research*, *164*(1–2), 69–76. [https://doi.org/10.1016/S0378-5955\(01\)00411-7](https://doi.org/10.1016/S0378-5955(01)00411-7)
- Klinge, A., & Klump, G. (2010). Mistuning detection and onset asynchrony in harmonic complexes in Mongolian gerbils. *The Journal of the Acoustical Society of America*, *128*(1), 280–290. <https://doi.org/10.1121/1.3436552>
- Klinge, A., & Klump, G. M. (2009). Frequency difference limens of pure tones and harmonics within complex stimuli in Mongolian gerbils and humans. *The Journal of the Acoustical Society of America*, *125*(1), 304–314. <https://doi.org/10.1121/1.3021315>
- Klink, K. B., & Klump, G. M. (2004). Duration discrimination in the mouse (*Mus musculus*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *190*(12). <https://doi.org/10.1007/s00359-004-0561-0>
- Kohlrausch, A., & Sander, A. (1995). Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets). *Journal of the Acoustical Society of America*, *97*(3), 1817–1829. <https://doi.org/10.1121/1.413097>
- Kohrman, D. C., Wan, G., Cassinotti, L., & Corfas, G. (2020). Hidden hearing loss: A disorder with multiple etiologies and mechanisms. *Cold Spring Harbor Perspectives in Medicine*, *10*(1), 1–20. <https://doi.org/10.1101/cshperspect.a035493>
- Köppl, C. (1997). Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*. *Journal of Neuroscience*, *17*(9), 3312–3321. <https://doi.org/10.1523/jneurosci.17-09-03312.1997>
- Kotak, V. C., Fujisawa, S., Lee, F. A., Karthikeyan, O., Aoki, C., & Sanes, D. H. (2005). Hearing loss raises excitability in the auditory cortex. *Journal of Neuroscience*, *25*(15). <https://doi.org/10.1523/JNEUROSCI.5169-04.2005>
- Krishna, B. S., & Semple, M. N. (2000). Auditory temporal processing: Responses to

- sinusoidally amplitude-modulated tones in the inferior colliculus. *Journal of Neurophysiology*, 84(1), 255–273. <https://doi.org/10.1152/jn.2000.84.1.255>
- Kujawa, S. G., & Liberman, M. C. (2006). Acceleration of age-related hearing loss by early noise exposure: Evidence of a misspent youth. *Journal of Neuroscience*, 26(7), 2115–2123. <https://doi.org/10.1523/JNEUROSCI.4985-05.2006>
- Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: Cochlear nerve degeneration after “temporary” noise-induced hearing loss. *Journal of Neuroscience*, 29(45), 14077–14085. <https://doi.org/10.1523/JNEUROSCI.2845-09.2009>
- Kujawa, S. G., & Liberman, M. C. (2015). Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss. *Hearing Research*, 330. <https://doi.org/10.1016/j.heares.2015.02.009>
- Kumar, U. A., Ameenudin, S., & Sangamanatha, A. V. (2012). Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. *Noise & Health*, 14(58), 100–105. <https://doi.org/10.4103/1463-1741.97252>
- Lauer, A. M., Dooling, R. J., Leek, M. R., & Poling, K. (2007). Detection and discrimination of simple and complex sounds by hearing-impaired Belgian Waterslager canaries. *The Journal of the Acoustical Society of America*, 122(6), 3615–3627. <https://doi.org/10.1121/1.2799482>
- Lauer, A. M., Molis, M., & Leek, M. R. (2009). Discrimination of time-reversed harmonic complexes by normal-hearing and hearing-impaired listeners. *JARO - Journal of the Association for Research in Otolaryngology*, 10(4), 609–619. <https://doi.org/10.1007/s10162-009-0182-y>
- Laumen, G., Tollin, D. J., Beutelmann, R., & Klump, G. M. (2016). Aging effects on the binaural interaction component of the auditory brainstem response in the Mongolian gerbil: Effects of interaural time and level differences. *Hearing Research*, 337, 46–58. <https://doi.org/10.1016/j.heares.2016.04.009>
- Leek, M., Dooling, R., Gleich, O., & Dent, M. L. (2005). Discrimination of temporal fine structure by birds and mammals. *Auditory Signal Processing, November 2015*, 470–476. [https://doi.org/10.1007/0-387-27045-0\\_57](https://doi.org/10.1007/0-387-27045-0_57)
- Leek, M. R., Dent, M. L., & Dooling, R. J. (2000). Masking by harmonic complexes in budgerigars ( *Melopsittacus undulatus* ) . *The Journal of the Acoustical Society of America*, 107(3). <https://doi.org/10.1121/1.428455>
- Liberman, C. (1982). Acute and chronic effects of acoustic trauma: Cochlear pathology and auditory nerve pathophysiology. *New Perspectives on Noise-Induced Hearing Loss.*, 105–134.
- Liberman, M. C. (1982). Single-neuron labeling in the cat auditory nerve. *Science*, 216(4551). <https://doi.org/10.1126/science.7079757>
- Liberman, M. Charles. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *Journal of the Acoustical Society of America*, 63(2), 442–455.

<https://doi.org/10.1121/1.381736>

- Liberman, M. Charles. (2017). Noise-induced and age-related hearing loss: New perspectives and potential therapies. In *F1000Research* (Vol. 6). <https://doi.org/10.12688/f1000research.11310.1>
- Liberman, M. Charles, & Dodds, L. W. (1984). Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hearing Research*, *16*(1). [https://doi.org/10.1016/0378-5955\(84\)90025-X](https://doi.org/10.1016/0378-5955(84)90025-X)
- Liberman, M. Charles, Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016). Toward a differential diagnosis of hidden hearing loss in humans. *PLoS ONE*, *11*(9), 1–15. <https://doi.org/10.1371/journal.pone.0162726>
- Liberman, M. Charles, & Kujawa, S. G. (2017). Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. In *Hearing Research* (Vol. 349). <https://doi.org/10.1016/j.heares.2017.01.003>
- Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B., & Ferrucci, L. (2011). Hearing loss and incident dementia. *Archives of Neurology*, *68*(2), 214–220. <https://doi.org/10.1001/archneurol.2010.362>
- Lin, H. W., Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2011). Primary neural degeneration in the guinea pig cochlea after reversible noise-induced threshold shift. *JARO - Journal of the Association for Research in Otolaryngology*, *12*(5). <https://doi.org/10.1007/s10162-011-0277-0>
- Lister, J. J., & Roberts, R. A. (2005). Effects of age and hearing loss on gap detection and the precedence effect: Narrow-band stimuli. *Journal of Speech, Language, and Hearing Research*, *48*(2). [https://doi.org/10.1044/1092-4388\(2005/033\)](https://doi.org/10.1044/1092-4388(2005/033))
- Lobarinas, E., Salvi, R., & Ding, D. (2013). Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hearing Research*, *302*. <https://doi.org/10.1016/j.heares.2013.03.012>
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18866–18869. <https://doi.org/10.1073/pnas.0607364103>
- Löscher, W., & Frey, H. H. (1987). One to three day dose intervals during subchronic treatment of epileptic gerbils with  $\gamma$ -vinyl GABA: anticonvulsant efficacy and alterations in regional brain GABA levels. *European Journal of Pharmacology*, *143*(3), 335–342. [https://doi.org/10.1016/0014-2999\(87\)90457-2](https://doi.org/10.1016/0014-2999(87)90457-2)
- Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *JARO - Journal of the Association for Research in Otolaryngology*, *12*(6). <https://doi.org/10.1007/s10162-011-0283-2>
- Marmel, F., Linley, D., Carlyon, R. P., Gockel, H. E., Hopkins, K., & Plack, C. J. (2013). Subcortical neural synchrony and absolute thresholds predict frequency

- discrimination independently. *JARO - Journal of the Association for Research in Otolaryngology*, 14(5). <https://doi.org/10.1007/s10162-013-0402-3>
- Marmel, Frederic, Plack, C. J., Hopkins, K., Carlyon, R. P., Gockel, H. E., & Moore, B. C. J. (2015). The role of excitation-pattern cues in the detection of frequency shifts in bandpass-filtered complex tones. *The Journal of the Acoustical Society of America*, 137(5), 2687–2697. <https://doi.org/10.1121/1.4919315>
- Mathew, A. K., Purdy, S. C., Welch, D., Pontoppidan, N. H., & Rønne, F. M. (2016). Electrophysiological and behavioural processing of complex acoustic cues. *Clinical Neurophysiology*, 127(1), 779–789. <https://doi.org/10.1016/j.clinph.2015.04.002>
- Mauermann, M., & Hohmann, V. (2007). Differences in loudness of positive and negative Schroeder-phase tone complexes as a function of the fundamental frequency. *The Journal of the Acoustical Society of America*, 121(2), 1028–1039. <https://doi.org/10.1121/1.2409772>
- Mener, D. J., Betz, J., Genther, D. J., Chen, D., & Lin, F. R. (2013). Hearing loss and depression in older adults. In *Journal of the American Geriatrics Society* (Vol. 61, Issue 9). <https://doi.org/10.1111/jgs.12429>
- Meyer, B. T., Jürgens, T., Wesker, T., Brand, T., & Kollmeier, B. (2010). Human phoneme recognition depending on speech-intrinsic variability. *The Journal of the Acoustical Society of America*, 128(5). <https://doi.org/10.1121/1.3493450>
- Miller, J. D. (1963). Deafening effects of noise on the cat. *Acta Otolaryngol (Stockh)*, 176, 1–91.
- Miller, R. L., Schilling, J. R., Franck, K. R., & Young, E. D. (1997). Effects of acoustic trauma on the representation of the vowel /ε/ in cat auditory nerve fibers. *The Journal of the Acoustical Society of America*, 101(6). <https://doi.org/10.1121/1.418321>
- Möhrle, D., Ni, K., Varakina, K., Bing, D., Lee, S. C., Zimmermann, U., Knipper, M., & Rüttiger, L. (2016). Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain. *Neurobiology of Aging*, 44, 173–184. <https://doi.org/10.1016/j.neurobiolaging.2016.05.001>
- Moore, B. C.J. (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America*, 54(3), 610–619. <https://doi.org/10.1121/1.1913640>
- Moore, Brian C. J., & Ernst, S. M. A. (2012). Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code. *The Journal of the Acoustical Society of America*, 132(3). <https://doi.org/10.1121/1.4739444>
- Moore, Brian C. J., Glasberg, B. R., Stoev, M., Füllgrabe, C., & Hopkins, K. (2012). The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). *The Journal of the Acoustical Society of America*, 131(2), 1003–1006. <https://doi.org/10.1121/1.3672808>
- Moore, Brian C. J., Hopkins, K., & Cuthbertson, S. (2009). Discrimination of complex

- tones with unresolved components using temporal fine structure information. *The Journal of the Acoustical Society of America*, 125(5), 3214.  
<https://doi.org/10.1121/1.3106135>
- Moore, Brian C. J., & Sek, A. (2009). Sensitivity of the human auditory system to temporal fine structure at high frequencies. *The Journal of the Acoustical Society of America*, 125(5), 3186. <https://doi.org/10.1121/1.3106525>
- Moore, Brian C. J., & Sek, A. (2011). Effect of level on the discrimination of harmonic and frequency-shifted complex tones at high frequencies. *The Journal of the Acoustical Society of America*, 129(5), 3206–3212.  
<https://doi.org/10.1121/1.3570958>
- Moore, Brian C.J. (2007). Cochlear Hearing Loss: Physiological, Psychological and Technical Issues: Second Edition. In *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues: Second Edition*.  
<https://doi.org/10.1002/9780470987889>
- Moore, Brian C.J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *JARO - Journal of the Association for Research in Otolaryngology*, 9(4), 399–406. <https://doi.org/10.1007/s10162-008-0143-x>
- Moore, Brian C.J. (2014). Auditory processing of temporal fine structure: Effects of age and hearing loss. In *Auditory Processing of Temporal Fine Structure: Effects of Age and Hearing Loss*. <https://doi.org/10.1142/9789814579667>
- Moore, Brian C.J., & Glasberg, B. R. (2004). A revised model of loudness perception applied to cochlear hearing loss. *Hearing Research*, 188(1–2).  
[https://doi.org/10.1016/S0378-5955\(03\)00347-2](https://doi.org/10.1016/S0378-5955(03)00347-2)
- Moore, Brian C.J., Glasberg, B. R., & Hopkins, K. (2006). Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hearing Research*, 222(1–2), 16–27.  
<https://doi.org/10.1016/j.heares.2006.08.007>
- Moore, Brian C.J., & Moore, G. A. (2003). Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hearing Research*, 182(1–2), 153–163.  
[https://doi.org/10.1016/S0378-5955\(03\)00191-6](https://doi.org/10.1016/S0378-5955(03)00191-6)
- Moore, Brian C.J., & Sek, A. (2009). Development of a fast method for determining sensitivity to temporal fine structure. *International Journal of Audiology*, 48(4), 161–171. <https://doi.org/10.1080/14992020802475235>
- Moore, Brian C.J., Vickers, D. A., & Mehta, A. (2012). The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *International Journal of Audiology*, 51(10), 715–721.  
<https://doi.org/10.3109/14992027.2012.690079>
- Moore, G. A., & Moore, B. C. J. (2003). Perception of the low pitch of frequency-shifted

- complexes. *The Journal of the Acoustical Society of America*, 113(2), 977–985.  
<https://doi.org/10.1121/1.1536631>
- Müller, M. (1996). The cochlear place-frequency map of the adult and developing mongolian gerbil. *Hearing Research*, 94(1–2), 148–156.  
[https://doi.org/10.1016/0378-5955\(95\)00230-8](https://doi.org/10.1016/0378-5955(95)00230-8)
- Neher, T., Lunner, T., Hopkins, K., & Moore, B. C. J. (2012). Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (L). *The Journal of the Acoustical Society of America*, 131(4).  
<https://doi.org/10.1121/1.3689850>
- Nelson, D. I., Nelson, R. Y., Concha-Barrientos, M., & Fingerhut, M. (2005). The global burden of occupational noise-induced hearing loss. *American Journal of Industrial Medicine*, 48(6), 446–458. <https://doi.org/10.1002/ajim.20223>
- Nelson, P. C., & Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America*, 116(4). <https://doi.org/10.1121/1.1784442>
- Nelson, P. C., & Carney, L. H. (2007). Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *Journal of Neurophysiology*, 97(1).  
<https://doi.org/10.1152/jn.00776.2006>
- Ohlemiller, K. K. (2009). Mechanisms and genes in human strial presbycusis from animal models. In *Brain Research* (Vol. 1277).  
<https://doi.org/10.1016/j.brainres.2009.02.079>
- Oxenham, A. J. (2016). Predicting the Perceptual Consequences of Hidden Hearing Loss. *Trends in Hearing*, 20, 1–6. <https://doi.org/10.1177/2331216516686768>
- Oxenham, A. J. (2018). How We Hear: The Perception and Neural Coding of Sound. In *Annual Review of Psychology* (Vol. 69). <https://doi.org/10.1146/annurev-psych-122216-011635>
- Oxenham, A. J., & Dau, T. (2001). Reconciling frequency selectivity and phase effects in masking. *The Journal of the Acoustical Society of America*, 110(3).  
<https://doi.org/10.1121/1.1394740>
- Oxenham, A. J., & Dau, T. (2004). Masker phase effects in normal-hearing and hearing-impaired listeners: Evidence for peripheral compression at low signal frequencies. *The Journal of the Acoustical Society of America*, 116(4), 2248–2257.  
<https://doi.org/10.1121/1.1786852>
- Özdamar, Ö., & Dallos, P. (1978). Synchronous responses of the primary auditory fibers to the onset of tone burst and their relation to compound action potentials. *Brain Research*, 155(1). [https://doi.org/10.1016/0006-8993\(78\)90320-7](https://doi.org/10.1016/0006-8993(78)90320-7)
- Pal, I., Paltati, C. R. B., Kaur, C., Shubhi Saini, Kumar, P., Jacob, T. G., Bhardwaj, D. N., & Roy, T. S. (2019). Morphological and neurochemical changes in GABAergic neurons of the aging human inferior colliculus. *Hearing Research*, 377.

<https://doi.org/10.1016/j.heares.2019.02.005>

- Palmer, A. R., & Moorjani, P. A. (1993). Responses to speech signals in the normal and pathological peripheral auditory system. *Progress in Brain Research*, 97(C).  
[https://doi.org/10.1016/S0079-6123\(08\)62268-2](https://doi.org/10.1016/S0079-6123(08)62268-2)
- Palmer, A. R., & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing Research*, 24(1). [https://doi.org/10.1016/0378-5955\(86\)90002-X](https://doi.org/10.1016/0378-5955(86)90002-X)
- Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. In *Annual Review of Neuroscience* (Vol. 21).  
<https://doi.org/10.1146/annurev.neuro.21.1.227>
- Parthasarathy, A., Hancock, K. E., Bennett, K., Degruittola, V., & Polley, D. B. (2020). Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *ELife*, 9.  
<https://doi.org/10.7554/eLife.51419>
- Peelle, J. E., & Wingfield, A. (2016). The Neural Consequences of Age-Related Hearing Loss. In *Trends in Neurosciences* (Vol. 39, Issue 7).  
<https://doi.org/10.1016/j.tins.2016.05.001>
- Peissig, J., & Kollmeier, B. (1997). Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *The Journal of the Acoustical Society of America*, 101(3).  
<https://doi.org/10.1121/1.418150>
- Pichora-Fuller, M. K., & Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *Journal of the Acoustical Society of America*, 91(4), 2129–2135.  
<https://doi.org/10.1121/1.403673>
- Pickles, J. O., Osborne, M. P., & Comis, S. D. (1987). Vulnerability of tip links between stereocilia to acoustic trauma in the guinea pig. *Hearing Research*, 25(2–3).  
[https://doi.org/10.1016/0378-5955\(87\)90089-X](https://doi.org/10.1016/0378-5955(87)90089-X)
- Pittman-Polletta, B. R., Wang, Y., Stanley, D. A., Schroeder, C. E., Whittington, M. A., & Kopell, N. J. (2020). 1 Differential contributions of synaptic 2 and intrinsic inhibitory currents to 3 speech segmentation via flexible 4 phase-locking in neural oscillators. *BioRxiv*, i, 1–59. <https://doi.org/10.1101/2020.01.11.902858>
- Plack, C. J., Léger, A., Prendergast, G., Kluk, K., Guest, H., & Munro, K. J. (2016). Toward a Diagnostic Test for Hidden Hearing Loss. *Trends in Hearing*, 20, 1–9.  
<https://doi.org/10.1177/2331216516657466>
- Plassmann, W., Peetz, W., & Schmidt, M. (1987). The cochlea in gerbilline rodents. *Brain, Behavior and Evolution*, 30(1–2). <https://doi.org/10.1159/000118639>
- Plomp, R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *Journal of the Acoustical Society of America*, 86(4).  
<https://doi.org/10.1121/1.398697>

- Prendergast, G., Couth, S., Millman, R. E., Guest, H., Kluk, K., Munro, K. J., & Plack, C. J. (2019). Effects of Age and Noise Exposure on Proxy Measures of Cochlear Synaptopathy. *Trends in Hearing*, *23*, 1–16.  
<https://doi.org/10.1177/2331216519877301>
- Prendergast, G., Guest, H., Munro, K. J., Kluk, K., Léger, A., Hall, D. A., Heinz, M. G., & Plack, C. J. (2017). Effects of noise exposure on young adults with normal audiograms I: Electrophysiology. *Hearing Research*, *344*, 68–81.  
<https://doi.org/10.1016/j.heares.2016.10.028>
- Prendergast, G., Millman, R. E., Guest, H., Munro, K. J., Kluk, K., Dewey, R. S., Hall, D. A., Heinz, M. G., & Plack, C. J. (2017). Effects of noise exposure on young adults with normal audiograms II: Behavioral measures. *Hearing Research*, *356*, 74–86. <https://doi.org/10.1016/j.heares.2017.10.007>
- Ralli, M., Greco, A., De Vincentiis, M., Sheppard, A., Cappelli, G., Neri, I., & Salvi, R. (2019). Tone-in-noise detection deficits in elderly patients with clinically normal hearing. *American Journal of Otolaryngology - Head and Neck Medicine and Surgery*, *40*(1). <https://doi.org/10.1016/j.amjoto.2018.09.012>
- Recio, A. (2001). Representation of harmonic complex stimuli in the ventral cochlear nucleus of the chinchilla. *The Journal of the Acoustical Society of America*, *110*(4), 2024–2033. <https://doi.org/10.1121/1.1397356>
- Recio, A., & Rhode, W. S. (2000). Basilar membrane responses to broadband stimuli. *The Journal of the Acoustical Society of America*, *108*(5), 2281–2298.  
<https://doi.org/10.1121/1.1318898>
- Rhode, W. S. (1971). Observations of the Vibration of the Basilar Membrane in Squirrel Monkeys using the Mössbauer Technique. *The Journal of the Acoustical Society of America*, *49*(4B). <https://doi.org/10.1121/1.1912485>
- Ridley, C. L., Kopun, J. G., Neely, S. T., Gorga, M. P., & Rasetshwane, D. M. (2018). Using thresholds in noise to identify hidden hearing loss in humans. *Ear and Hearing*, *39*(5). <https://doi.org/10.1097/AUD.0000000000000543>
- Riedel, H., Granzow, M., & Kollmeier, B. (2001). Single-sweep-based methods to improve the quality of auditory brain stem responses Part II: Averaging methods. *Zeitschrift Fur Audiologie*, *40*(2), 62–85.
- Robert Frisina, D., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research*, *106*(1–2). [https://doi.org/10.1016/S0378-5955\(97\)00006-3](https://doi.org/10.1016/S0378-5955(97)00006-3)
- Robinson, D. W., & Sutton, G. J. (1979). Age effect in hearing - a comparative analysis of published threshold data. *International Journal of Audiology*, *18*(4), 320–334.  
<https://doi.org/10.1080/00206097909072634>
- Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. In *Physiological Reviews* (Vol. 81, Issue 3).  
<https://doi.org/10.1152/physrev.2001.81.3.1305>



- Rogalla, M. M., & Jannis Hildebrandt, K. (2020). Aging but not age-related hearing loss dominates the decrease of parvalbumin immunoreactivity in the primary auditory cortex of mice. *ENeuro*, 7(3). <https://doi.org/10.1523/ENEURO.0511-19.2020>
- Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, 30(4). <https://doi.org/10.1152/jn.1967.30.4.769>
- Rosen, M. J., Sarro, E. C., Kelly, J. B., & Sanes, D. H. (2012). Diminished behavioral and neural sensitivity to sound modulation is associated with moderate developmental hearing loss. *PLoS ONE*, 7(7). <https://doi.org/10.1371/journal.pone.0041514>
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. In *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* (Vol. 336, Issue 1278). <https://doi.org/10.1098/rstb.1992.0070>
- Ruggero, M. A., & Rich, N. C. (1991). Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane. *Journal of Neuroscience*, 11(4). <https://doi.org/10.1523/jneurosci.11-04-01057.1991>
- Ruggero, M. A., Robles, L., & Rich, N. C. (1992). Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *Journal of Neurophysiology*, 68(4). <https://doi.org/10.1152/jn.1992.68.4.1087>
- Ruggero, Mario A. (1992). Responses to sound of the basilar membrane of the mammalian cochlea. *Current Opinion in Neurobiology*, 2(4). [https://doi.org/10.1016/0959-4388\(92\)90179-O](https://doi.org/10.1016/0959-4388(92)90179-O)
- Ruggero, Mario A., Narayan, S. S., Temchin, A. N., & Recio, A. (2000). Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: Comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), 11744–11750. <https://doi.org/10.1073/pnas.97.22.11744>
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology*, 22(15). <https://doi.org/10.1016/j.cub.2012.05.025>
- Ryan, A. (1976). Hearing sensitivity of the mongolian gerbil, *Meriones unguiculatus*. *Journal of the Acoustical Society of America*, 59(5), 1222–1226. <https://doi.org/10.1121/1.380961>
- Sachs, M. B., & Abbas, P. J. (1974). Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli. *Journal of the Acoustical Society of America*, 56(6), 1835–1847. <https://doi.org/10.1121/1.1903521>
- Schaette, R., & McAlpine, D. (2011). Tinnitus with a normal audiogram: Physiological evidence for hidden hearing loss and computational model. *Journal of Neuroscience*, 31(38), 13452–13457. <https://doi.org/10.1523/JNEUROSCI.2156->

11.2011

- Schalk, T. B., & Sachs, M. B. (1980). Nonlinearities in auditory-nerve fiber responses to bandlimited noise. *Journal of the Acoustical Society of America*, 67(3).  
<https://doi.org/10.1121/1.383970>
- Schmiedt, R. A., Mills, J. H., & Boettcher, F. A. (1996). Age-related loss of activity of auditory-nerve fibers. *Journal of Neurophysiology*, 76(4), 2799–2803.  
<https://doi.org/10.1152/jn.1996.76.4.2799>
- Schmiedt, Richard A. (1989). Spontaneous rates, thresholds and tuning of auditory-nerve fibers in the gerbil: Comparisons to cat data. *Hearing Research*, 42(1), 23–35. [https://doi.org/10.1016/0378-5955\(89\)90115-9](https://doi.org/10.1016/0378-5955(89)90115-9)
- Schmiedt, Richard A. (1996). Effects of aging on potassium homeostasis and the endocochlear potential in the gerbil cochlea. *Hearing Research*, 102(1–2).  
[https://doi.org/10.1016/S0378-5955\(96\)00154-2](https://doi.org/10.1016/S0378-5955(96)00154-2)
- Schmiedt, Richard A. (2010). *The Physiology of Cochlear Presbycusis*.  
[https://doi.org/10.1007/978-1-4419-0993-0\\_2](https://doi.org/10.1007/978-1-4419-0993-0_2)
- Schomann, T., Ramekers, D., De Groot, J. C. M. J., Van Der Ploeg, C. H., Hendriksen, F. G. J., Böhringer, S., Klis, S. F. L., Frijns, J. H. M., & Huisman, M. A. (2018). Ouabain Does Not Induce Selective Spiral Ganglion Cell Degeneration in Guinea Pigs. *BioMed Research International*, 2018.  
<https://doi.org/10.1155/2018/1568414>
- Schreiner, C. E., & Langner, G. (1988). Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *Journal of Neurophysiology*, 60(6).  
<https://doi.org/10.1152/jn.1988.60.6.1823>
- Schroeder, M. R. (1970). Synthesis of Low-Peak-Factor Signals and Binary Sequences with Low Autocorrelation. *IEEE Transactions on Information Theory*, 16(1).  
<https://doi.org/10.1109/TIT.1970.1054411>
- Schuknecht, H. F., & Gacek, M. R. (1993). Cochlear pathology in presbycusis. *Annals of Otolaryngology, Rhinology and Laryngology*, 102(1 II).  
<https://doi.org/10.1177/00034894931020s101>
- Schulte, B. A., & Schmiedt, R. A. (1992). Lateral wall Na, K-ATPase and endocochlear potentials decline with age in quiet-reared gerbils. *Hearing Research*, 61(1–2), 35–46. [https://doi.org/10.1016/0378-5955\(92\)90034-K](https://doi.org/10.1016/0378-5955(92)90034-K)
- Schulze, H., & Langner, G. (1997). Periodicity coding in the primary auditory cortex of the Mongolian gerbil (*Meriones unguiculatus*): Two different coding strategies for pitch and rhythm? *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*, 181(6), 651–663. <https://doi.org/10.1007/s003590050147>
- Seligmann, H., Podoshin, L., Ben-David, J., Fradis, M., & Goldsher, M. (1996). Drug-induced tinnitus and other hearing disorders. *Drug Safety*, 14(3), 198–212.  
<https://doi.org/10.2165/00002018-199614030-00006>
- Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1982). Measurement of basilar

- membrane motion in the guinea pig using the Mössbauer technique. *Journal of the Acoustical Society of America*, 72(1). <https://doi.org/10.1121/1.387996>
- Sergeyenko, Y., Lall, K., Charles Liberman, M., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: An early-onset contributor to auditory functional decline. *Journal of Neuroscience*, 33(34), 13686–13694. <https://doi.org/10.1523/JNEUROSCI.1783-13.2013>
- Shamma, S. A. (1985). Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve. *Journal of the Acoustical Society of America*, 78(5), 1622–1632. <https://doi.org/10.1121/1.392800>
- Sinex, D. G., López, D. E., & Warr, W. B. (2001). Electrophysiological responses of cochlear root neurons. *Hearing Research*, 158(1–2). [https://doi.org/10.1016/S0378-5955\(01\)00293-3](https://doi.org/10.1016/S0378-5955(01)00293-3)
- Sinnott, J. M., Brown, C. H., & Brown, F. E. (1992). Frequency and intensity discrimination in Mongolian gerbils, African monkeys and humans. *Hearing Research*, 59(2). [https://doi.org/10.1016/0378-5955\(92\)90117-6](https://doi.org/10.1016/0378-5955(92)90117-6)
- Sinnott, J. M., & Mosqueda, S. B. (2003). Effects of aging on speech sound discrimination in the mongolian gerbil. *Ear and Hearing*, 24(1), 30–37. <https://doi.org/10.1097/01.AUD.0000051747.58107.89>
- Slama, M. C. C., Ravicz, M. E., & Rosowski, J. J. (2010). Middle ear function and cochlear input impedance in chinchilla. *The Journal of the Acoustical Society of America*, 127(3), 1397–1410. <https://doi.org/10.1121/1.3279830>
- Smith, B. K., Sieben, U. K., Kohlrausch, A., & Schroeder, M. R. (1986). Phase effects in masking related to dispersion in the inner ear. *Journal of the Acoustical Society of America*, 80(6), 1631–1637. <https://doi.org/10.1121/1.394327>
- Snell, K. B. (1997). Age-related changes in temporal gap detection. *The Journal of the Acoustical Society of America*, 101(4), 2214–2220. <https://doi.org/10.1121/1.418205>
- Sommers, M. S., & Gehr, S. E. (1998). Auditory suppression and frequency selectivity in older and younger adults. *The Journal of the Acoustical Society of America*, 103(2). <https://doi.org/10.1121/1.421220>
- Spankovich, C., Le Prell, C. G., Lobarinas, E., & Hood, L. J. (2017). Noise history and auditory function in young adults with and without type 1 diabetes mellitus. *Ear and Hearing*, 38(6). <https://doi.org/10.1097/AUD.0000000000000457>
- Spoendlin, H. (1971). Primary structural changes in the organ of corti after acoustic overstimulation. *Acta Oto-Laryngologica*, 71(1–6). <https://doi.org/10.3109/00016487109125346>
- Stamper, G. C., & Johnson, T. A. (2015). Auditory function in normal-hearing, noise-exposed human ears. *Ear and Hearing*, 36(2). <https://doi.org/10.1097/AUD.0000000000000107>

- Strouse, A., Ashmead, D. H., Ohde, R. N., & Grantham, D. W. (1998). Temporal processing in the aging auditory system. *The Journal of the Acoustical Society of America*, *104*(4). <https://doi.org/10.1121/1.423748>
- Stuart, A., & Phillips, D. P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. *Ear and Hearing*, *17*(6). <https://doi.org/10.1097/00003446-199612000-00004>
- Summers, V., de Boer, E., & Nuttall, A. L. (2003). Basilar-membrane responses to multicomponent (Schroeder-phase) signals: Understanding intensity effects. *The Journal of the Acoustical Society of America*, *114*(1), 294–306. <https://doi.org/10.1121/1.1580813>
- Tarnowski, B. I., Schmiedt, R. A., Hellstrom, L. I., Lee, F. S., & Adams, J. C. (1991). Age-related changes in cochleas of mongolian gerbils. *Hearing Research*, *54*(1). [https://doi.org/10.1016/0378-5955\(91\)90142-V](https://doi.org/10.1016/0378-5955(91)90142-V)
- Temchin, A. N., & Ruggero, M. A. (2010). Phase-locked responses to tones of chinchilla auditory nerve fibers: Implications for apical cochlear mechanics. *JARO - Journal of the Association for Research in Otolaryngology*, *11*(2). <https://doi.org/10.1007/s10162-009-0197-4>
- Valero, M. D., Burton, J. A., Hauser, S. N., Hackett, T. A., Ramachandran, R., & Liberman, M. C. (2017). Noise-induced cochlear synaptopathy in rhesus monkeys (*Macaca mulatta*). *Hearing Research*, *353*. <https://doi.org/10.1016/j.heares.2017.07.003>
- Van Rossum, M. C. W. (2001). A novel spike distance. *Neural Computation*, *13*(4). <https://doi.org/10.1162/089976601300014321>
- Verschooten, E., Robles, L., Desloovere, C., & Joris, P. (2013). Assessment of neural phase-locking at the round window in human. *Assoc Res Otolaryngol Abs*, *426*.
- Verschooten, Eric, Shamma, S., Oxenham, A. J., Moore, B. C. J., Joris, P. X., Heinz, M. G., & Plack, C. J. (2019). The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints. *Hearing Research*, *377*, 109–121. <https://doi.org/10.1016/j.heares.2019.03.011>
- Versfeld, N. J., & Dreschler, W. A. (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *The Journal of the Acoustical Society of America*, *111*(1). <https://doi.org/10.1121/1.1426376>
- Versteegh, C. P. C., Meenderink, S. W. F., & Van Der Heijden, M. (2011). Response characteristics in the apex of the gerbil cochlea studied through auditory nerve recordings. *JARO - Journal of the Association for Research in Otolaryngology*, *12*(3), 301–316. <https://doi.org/10.1007/s10162-010-0255-y>
- Von Békésy, G., & Wever, E. G. (1960). *Experiments in hearing* (Vol. 195). McGraw-Hill New York.

- von Helmholtz, H. (1877). Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik. In *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik*. <https://doi.org/10.1007/978-3-663-18653-3>
- Walton, J. P. (2010). Timing is everything: Temporal processing deficits in the aged auditory brainstem. *Hearing Research*, *264*(1–2). <https://doi.org/10.1016/j.heares.2010.03.002>
- Wever, E. G., & Bray, C. W. (1937). The perception of low tones and the resonance-volley theory. *Journal of Psychology: Interdisciplinary and Applied*, *3*(1). <https://doi.org/10.1080/00223980.1937.9917483>
- Winter, I. M., & Palmer, A. R. (1991). Intensity coding in low-frequency auditory-nerve fibers of the guinea pig. *Journal of the Acoustical Society of America*, *90*(4). <https://doi.org/10.1121/1.401675>
- Woolf, N. K., Ryan, A. F., & Bone, R. C. (1981). Neural phase-locking properties in the absence of cochlear outer hair cells. *Hearing Research*, *4*(3–4). [https://doi.org/10.1016/0378-5955\(81\)90017-4](https://doi.org/10.1016/0378-5955(81)90017-4)
- World Health Organisation. (2018). *Ageing and health*. <https://www.who.int/newsroom/fact-sheets/detail/ageing-and-health>
- World Health Organisation. (2020). *Deafness and hearing loss*. <https://www.who.int/newsroom/fact-sheets/detail/deafness-and-hearing-loss>
- Wright, A., Davis, A., Bredberg, G., Ulehlova, L., & Spencer, H. (1987). Hair cell distributions in the normal human cochlea. *Acta Oto-Laryngologica. Supplementum*, *444*.
- Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., & Liberman, M. C. (2019). Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear. *Neuroscience*, *407*, 8–20. <https://doi.org/10.1016/j.neuroscience.2018.07.053>
- Wu, Pei Zhe, O'Malley, J. T., de Gruttola, V., & Charles Liberman, M. (2020). Age-related hearing loss is dominated by damage to inner ear sensory cells, not the cellular battery that powers them. *Journal of Neuroscience*, *40*(33), 6357–6366. <https://doi.org/10.1523/JNEUROSCI.093720.2020>
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46. [https://doi.org/10.1016/S1364-6613\(00\)01816-7](https://doi.org/10.1016/S1364-6613(00)01816-7)
- Zhang, L., Engler, S., Koepcke, L., Steenken, F., & Köppl, C. (2018). Concurrent gradients of ribbon volume and AMPA-receptor patch volume in cochlear afferent synapses on gerbil inner hair cells. *Hearing Research*, *364*, 81–89. <https://doi.org/10.1016/j.heares.2018.03.028>

## 6.2 List of Figures

Figure 1: Schematic representation of phase-locked ANFs in the response of an amplitude modulated pure tone.....	6
Figure 2: Temporal waveforms of SCHR complexes .....	22
Figure 3: Alignment of the spike trains elicited by SCHR+ and SCHR- stimuli prior to the vR analysis.....	28
Figure 4: Example responses of two different auditory-nerve fibers.....	34
Figure 5: Example responses of four different neurons.....	36
Figure 6: Sensitivity $d'$ (mean, $\pm$ SE) for discriminating a SCHR complex target stimulus of opposite-sign $C$ value.....	38
Figure 7: Sensitivity $d'$ (mean, $\pm$ SE) for discriminating a SCHR complex target stimulus from a SCHR complex reference stimulus with $C = 1$ .....	43
Figure 8: Sensitivity $d'$ (mean, $\pm$ SE) for discriminating a SCHR complex target stimulus from a SCHR complex reference stimulus for the reference $C = -1$ .....	44
Figure 9: Sensitivity $d'$ (mean, $\pm$ SE) for discriminating a SCHR complex target stimulus from a SCHR complex reference stimulus for the reference $C = 0$ .....	45
Figure 10: Comparison of mean behavioral discrimination sensitivity, $d'$ , with neuronal pooled sensitivity, $d'_p$ , obtained in the sweep-direction experiment .....	50
Figure 11: Comparison of mean behavioral discrimination sensitivity $d'$ with neuronal pooled sensitivity $d'_p$ obtained in the sweep-velocity experiment.....	53
Figure 12: Temporal waveforms of Schroeder phase (SCHR) complexes with different values for fundamental frequency $f_0$ and $C$ .....	65
Figure 13: ABR wave I amplitude (mean $\pm$ SE) of the three old gerbils shown for each ear separately (grey and black). The solid lines indicate the linear regressions.....	68
Figure 14: Mean sensitivity $d'$ with SE in relation to absolute target $C$ .....	70

Figure 15: Temporal waveforms of harmonic (H) and inharmonic (I) tone complexes in the TFS1 test..... 84

Figure 16: Power spectra of H (blue) and I (red) complexes with frequency shift  $\Delta f = 42.5\%$  of  $f_0$  embedded in pink noise..... 85

Figure 17: Effect of click level in dB SPL on wave I amplitude (mean  $\pm$  SE) in  $\mu\text{V}$  for both ears combined..... 91

Figure 18: Sensitivity  $d'$  (mean  $\pm$  SE) in relation to frequency shift  $\Delta f$  between the harmonic and inharmonic complex. .... 93

Figure 19: Comparison of relative TFS1 thresholds ( $\Delta f * f_0 / f_c$ ,  $\Delta f$ ) for gerbils of the different groups and conditions with performance of humans..... 98

### 6.3 List of Tables

Table 1: ABR threshold in response to click stimuli in dB SPL for left and right ear with the indication of group, age and treatment for all tested gerbils..... 90

Table 2: Mean  $\Delta f$  threshold ( $d' = 1$ ) in relation to condition and group..... 94

## 6.4 Declaration

Herewith I declare that the dissertation has, neither as a whole, nor in part, been submitted for assessment in a doctoral procedure at another university. Furthermore, I confirm that I have written this dissertation myself an independently, using only the indicated references and resources. My work as a doctoral student has always been performed under consideration of the guidelines for “good scientific practice” of the Carl v. Ossietzky University Oldenburg.

### Erklärung

Hiermit bestätige ich, dass diese Dissertation, weder in ihrer Gesamtheit noch in Teilen, einer anderen wissenschaftlichen Hochschule zur Begutachtung in einem Promotionsverfahren vorgelegen hat. Des Weiteren erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegeben Literatur und Hilfsmittel angefertigt habe. Meine Arbeit als Doktorand wurde immer unter Berücksichtigung der Leitlinien für „gute wissenschaftliche Praxis“ durchgeführt.

Oldenburg, 31.03.2021\_\_\_\_\_



## 6.5 Danksagung

Viele Menschen haben mir sehr während meiner Promotion geholfen.

Als Erstes möchte ich Susanne Groß danken, der besten technischen Mitarbeiterin überhaupt. Ich bin sehr glücklich, dass ich mit ihr zusammenarbeiten konnte. Vielen Dank für die Hilfe bei den Versuchen und für die Betreuung der Gerbils. Ich konnte mich stets blind auf ihrer Arbeit verlassen.

Georg Klump danke ich für die Möglichkeit, diese Promotion durchzuführen und die langjährige Betreuung. Ich bedanke mich außerdem für die Möglichkeit an internationalen Tagungen teilzunehmen und für die große Hilfe an den Manuskripten.

Christine Köppl danke ich für erweiterte Betreuung und für die Begutachtung der Disputation.

Peter Heil danke ich für die Übernahme des Zweitgutachtens.

Meike Rogalla danke ich für die große Unterstützung bei allen Problemen, die zum Ende meiner Promotion aufkamen. Außerdem danke ich für die schöne Zeit im Turmzimmer. Vielen Dank!

Katja Bleckman danke ich die nette Zerstreung während der Kaffeepausen und der Hilfe bei allen anderen Dingen.

Rainer Beutelmann danke ich für die große Hilfe bei technischen Aspekten. Ohne ihn wäre jeder Doktorand verloren.

Friederike Könitz-Steenken danke ich für die Hilfe bei den ouabain Gerbils. Danke, dass du meiner Fehler immer so locker genommen hast.

Außerdem danke ich meiner Arbeitsgruppe für die schöne Zeit.

Frank Magerquark danke ich für die Hilfe zum Ende meiner Promotion.

Zuletzt möchte ich meinen Eltern danken für die Unterstützung während des Studiums. Sollte ich mal Kinder haben, will ich so sein wie ihr.