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

Friederike Steenken1 | Henning Oetjen1 | Rainer Beutelmann1 Ⓡ | Laurel H. Carney2,3 Ⓡ | Christine Koeppl1 Ⓡ | Georg M. Klump1 Ⓡ

1Cluster of Excellence “Hearing4all” and Research Centre Neurosensory Science, Department of Neuroscience, School of Medicine and Health Science, Carl von Ossietzky University Oldenburg, Oldenburg, Germany
2Department of Biomedical Engineering, University of Rochester, Rochester, New York, USA
3Hanse-Wissenschaftskolleg, Delmenhorst, Germany

Correspondence
Georg Klump, Cluster of Excellence “Hearing4all” and Research Centre Neurosensory Science, Department of Neuroscience, School of Medicine and Health Science, Carl von Ossietzky University, Oldenburg, 26111 Oldenburg, Germany.
Email: georg.klump@uol.de

Funding information
Hanse-Wissenschaftskolleg; National Institutes of Health, Grant/Award Number: NIH-R01-001641; Deutsche Forschungsgemeinschaft

Abstract
Schroeder-phase harmonic tone complexes have been used in physiological and psychophysical studies in several species to gain insight into cochlear 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 Mongolian 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 of either 1 ms or related to the stimulus’ period, was compared with average discharge rates. Neural responses and behavioral performance were both expressed in terms of sensitivity, $d’$, to allow direct comparisons. Our results suggest that in the auditory nerve, stimulus fine structure is represented by spike timing, whereas envelope is represented by rate. In the inferior colliculus, both temporal fine structure and envelope appear to be represented best by rate. 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. Furthermore, the high sensitivity observed in the inferior colliculus neural rate-based discrimination suggests that these neurons integrate across multiple inputs arising from the auditory periphery.

KEYWORDS
auditory nerve, inferior colliculus, perception, temporal envelope, temporal fine structure
1 | INTRODUCTION

Schroeder-phase (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 cochlear function and neural coding, including the phase response of auditory filters (e.g., Carlyon et al., 2017; Kohlrausch & Sander, 1995; Oxenham & Dau, 2001, 2004), the effect of the phase spectrum on masking (e.g., Carlyon & Datta, 1997; Smith et al., 1986), and sensitivity to temporal fine structure (TFS; e.g., Dooling et al., 2002; Drennan et al., 2008).

SCHR complexes can be manipulated by varying the phase spectrum, without changing the magnitude spectrum. The fast linear frequency sweeps of SCHR complexes occur within each fundamental period of the stimulus; the sweeps are either upward (SCHR+/C0) or downward (SCHR−), depending on the sign of the phase spectrum (Figure 1). 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 magnitude spectra and similar acoustical envelopes (ENVs). To date, the perception of SCHR complexes has been evaluated primarily in birds (Dooling et al., 2002) and humans (Carlyon et al., 2017; Drennan et al., 2008; Kohlrausch & Sander, 1995; Oxenham & Dau, 2001, 2004), whereas neural studies have focused on small mammals (e.g., Cedolin & Delgutte, 2010; Recio, 2001). 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.

Cochlear mechanical measurements in the base confirm that SCHR+ complexes elicit peakier patterns of excitation on the basilar membrane (BM), as compared with SCHR− complexes (chinchilla: Recio & Rhode, 2000; guinea pig: Summers et al., 2003), consistent with a high-frequency-based cochlear model for the masking differences of SCHR stimuli (Kohlrausch & Sander, 1995). Peakier SCHR+ responses are more vulnerable to cochlear compression, consistent with reduced responses to and weaker masking of SCHR+ stimuli, emphasizing distinct TFS and 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 cochlear nucleus in the chinchilla (Recio, 2001) and responses to SCHR− complexes in AN fibers of the cat (Cedolin & Delgutte, 2010). Recio (2001) reported higher firing rates for SCHR− than for SCHR+ complexes for neurons with high characteristic frequencies (CFs; the frequency to which a neuron is most sensitive), that is, CFs > 3–4 kHz, whereas for neurons with CFs below 2 kHz, the difference between responses to SCHR+ and SCHR− was diminished. The available BM and neural results in animal models deviate somewhat from psychophysical findings in humans: Threshold differences between SCHR+ and SCHR− maskers in humans are about 20 dB for a signal frequency of 1 kHz, and at a signal frequency of 250 Hz, the difference still

---

**Figure 1** Temporal waveforms of SCHR complexes, with different values for the fundamental frequency (f0) and the parameter C, which determines the speed of the frequency sweep and, therefore, the duty cycle (absolute value of parameter C). Gray lines represent the waveform, dark lines the ENV. 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 f0 on the waveform.
reaches 6 dB (Oxenham & Dau, 2001). The discrepancy between physiological and psychophysical results may be attributed to different species-specific apical–basal transition points in the cochlear mechanics (Shera et al., 2010). To address this discrepancy, we must bridge the gap in our understanding of responses of low-frequency neurons to low-frequency SCHR complexes in mammals. Our understanding is also benefited by directly comparing perceptual and neuronal sensitivity for discrimination of SCHR complexes within the same species.

In this study, we tested whether temporal or average-discharge-rate information available at the single-unit level was sufficient to explain behavioral discrimination. 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 inferior colliculus (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.

We presented gerbils with similar 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 for a deeper understanding of the processing of harmonic tone complexes and frequency sweeps.

2 | MATERIALS AND METHODS

2.1 | 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 five gerbils (2 M, 3 F, age 7–8 months) and in the central nucleus of the IC in 11 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 with 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, auditory brainstem responses (ABR) 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 experiments and 70 mg/kg for ABR measurements prior to behavioral experiments) and xylazine (6 mg/kg for AN and IC experiments and 3 mg/kg for ABR measurements prior to behavioral experiments) 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 experiments and behavior; by 10^3 for AN experiments) and band-pass filtered (0.3–3 kHz) using an ISO-80 preamplifier (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 et al. (2016). During neurophysiological experiments, ABR thresholds were rechecked periodically, and the experiment was terminated if the 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.

2.2 Schroeder-phase-complex stimuli

SCHR complexes used in the behavioral and AN experiments were equal-amplitude harmonic-tone complexes with frequencies ranging from fundamental frequency ($f_0$) 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 $n$th harmonic, $n$ was the harmonic number, and $N$ was the total number of harmonics (e.g., 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 (Figure 1). 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 sweep direction and sweep velocity 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 ENVs and long-term spectra are similar (compare examples in Figure 1a,b). In the sweep-velocity experiment, the reference was a SCHR complex with a $C$ value of $-1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$, and offset. The overall stimulus level for all SCHR stimuli was 60 dB SPL.

2.3 AN recordings

Gerbils were anesthetized intraperitoneally (i.p.) with a combination of ketamine (135 mg/kg) and xylazine (6 mg/kg) 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 nonsteroidal 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 ear bar at the bony edge of the ear canal for delivery of the sound. The ear bar 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 body weight, 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 $\sim$10–30 MΩ, 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.), with 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 nonsteroidal 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 ear bar at the bony edge of the ear canal for delivery of the sound. The ear bar 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 body weight, 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 $\sim$10–30 MΩ, 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 ear bar sealed to the ear canal. The sound system was individually calibrated at the start of each experiment via a microphone (ER7C, Etymotic Research Inc., Elk Grove Village, IL, USA) that was integrated into the ear bar, using a microphone amplifier (MA3, TDT Inc.).

Noise search stimuli (1–9 kHz at 50 dB SPL, 50-ms duration) were played, whereas the electrode was moved forward into the nerve. When a unit was isolated, its frequency response range was first assessed audiodesiically using 50-ms duration tones. The best frequency (BF) of each fiber was then determined with tones of 50-ms duration, including 5-ms cosine rise/fall times, at a fixed level of 0–30 dB above the audiodesiically assessed threshold within approximately ±1 kHz around the audiodesiically 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 f0 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 band-pass 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 cochlear nucleus: (1) presence of a prepotential in the spike waveform, characteristic for anteroventral cochlear nucleus (Keine & Ruebsamen, 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).

2.4 IC recordings

Gerbils were anesthetized with an initial i.p. injection of ketamine (135 mg/kg body weight) and xylazine hydrochloride (6 mg/kg body weight). A surgical plane of anesthesia was maintained with subcutaneous injections (0.03–0.05 mL) or slow subcutaneous 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 μL/h. Subcutaneous 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 centimeters 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 MΩ-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.). The neural signals were amplified using an ISO-80 preamplifier (World Precision Instruments) and sampled using a Hammerfall digital signal processor (DSP, Multiface II, RME Audio; sampling frequency 48 kHz) controlled by custom MATLAB software. 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-cosine ramps) presented every 600 ms, with levels spanning 10–70 dB SPL and frequencies in a 2-octave range centered on an audiovisual estimate of CF. Neurons were further characterized by a modulation transfer function (MTF) using 500-ms duration, sinusoidentally amplitude-modulated tones at CF, with modulation frequencies from 4 to 512 Hz and five 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 f0 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 f0s 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 μA 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 (Schnupp et al., 2015), 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.

2.5 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 (vR) distance, which provides a metric for the difference between two spike trains, computed for a given temporal precision (van Rossum, 2001). The vR 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 (τ) 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; neurons are most likely to respond 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 Figure 2a,b). Direct comparisons of spike trains, without compensation for the different position of BF within the instantaneous-frequency sweep, can lead to large vR 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 (Figure 2c). Lags larger than one f0 period were wrapped back into the single period range. vR distances were then computed using the temporally aligned responses.

To calculate the vR distances, one group of spike trains was defined as the background or reference condition, and 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 vR analysis were τ = 1 ms and τ = 1/f0. The integration time constant of 1/f0 matched the period in the ENV of the SCHR stimuli. Thus, the vR integration times used were 1 ms (for all stimuli), and 2.5, 5, 10, and 25 ms for each of the f0s tested. Because the resolution of the recorded spike times was 20 μs, and the exponential decay for the vR analysis started at the exact spike times, a time constant of 1 ms in the vR analysis was suitable to evaluate the neural representation of the stimulus 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
Alignment of the spike trains elicited by SCHR+ and SCHR− stimuli prior to the vR analysis. (a) Raster plots of the response of an AN fiber (BF 1250 Hz, mean spontaneous rate of <1 spikes per second) stimulated with a 400-ms SCHR complex with a f0 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 vR 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 vR analysis.

A d’ higher than 1 represented a discriminable vR or rate difference (Green & Swets, 1966). Values of d’ were limited to 4.3, because they were calculated based on 30 repetitions of each stimulus condition. The effects of the main factors (e.g., C value, f0) 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 errors. To compare the neural sensitivities with the behavioral outcomes, the sensitivities of individual neurons were optimally combined across pairs of neurons, assuming that on average, the total information of two neurons was sufficient to explain behavioral sensitivity. The displayed values of d’p are the grand mean across all pairwise combinations, N being the total number of neurons, excluding self-comparisons:

\[ d’_p = \frac{2}{N(N-1)} \sum_{n=1}^{N} \sum_{m=n+1}^{N} \sqrt{d^2_n + d^2_m} \]

Pooling more than two neurons in the same way increased the value of d’p, but did not lead to overall higher correlations between neural and behavioral sensitivities.

### Behavioral experiments

#### 2.6.1 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 $T_{30}$ 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 soundcard (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 to 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.

### 2.6.2 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, approximately 75% of the trials in a session) or equal to the reference stimulus (sham trial, approximately 25% of the trials in a session). 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 reentering 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 an expected easy discrimination (i.e., $f_0 = 50$ Hz, $C = \pm 1$). 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 stable stimulus control in three consecutive sessions. Training typically required about 150 sessions, including the initial training of the waiting period on the pedestal. Thus, at the beginning of data collection for each experiment, the animals were well trained, independent of their previous experience. Stability of thresholds over the total testing period was confirmed by comparing the first and the last threshold estimates.

Each test session started with a warm-up 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.

2.6.3 | 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.

2.7 | 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.

2.8 | 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 $f_0$ as fixed factors. The results of these analyses are presented above each subpanel of Figures 5–8. Subsequent additional GLMM ANOVAs included more factors, for example, a classification of the BF for AN and IC units into sets with BF above or below 1.850 kHz. In the sweep-direction experiment, 10 and 12 AN fibers had a BF below or above 1.850 kHz, respectively, and in the sweep-velocity experiment, 8 and 12 AN fibers had a BF below or above 1.850 kHz, respectively. In the direction and the sweep-velocity experiment, 24 and 21 IC units had a BF below or above 1.850 kHz, respectively. For the IC, the MTF was classified into either band-pass (17 units), band-reject (16 units), or hybrid (12 units) MTF type, and the BMF was classified into BMF classes of up to 75 Hz (11 units), between 75 and 150 Hz (28 units), or 150 Hz and above (6 units). Only main effects from the GLMM ANOVA results are reported below. Supplementary material provides the results of the GLMM ANOVAs with all factors that were measured in AN and IC including full interactions (Supporting Information).

In addition to the analysis of the relation between sensitivity and the parameters characterizing the SCHR-stimulus waveform, we also conducted an analysis relating sensitivity to a measure of differences in the speed and direction of instantaneous frequency along the frequency map of the cochlea (see Supporting Information). Using the parameters $C$ value, $f_0$, and $N$ in the expression defining the SCHR-stimulus waveform, we calculated the time it would take the frequency sweep to traverse 1 μm of a gerbil’s BM. This value is directly proportional to the time it would take to traverse one critical band (Kittel et al., 2002). Thus, differences in this measure between reference and target stimulus may reflect differences in the transition of the pattern of excitation along the cochlear map. Because this measure reflecting the sweep speed and direction in a physiology-inspired dimension is highly correlated with the parameters $C$ value, $f_0$, and $N$, we decided to present the data in the results section in relation to parameters $C$ value and $f_0$ avoiding redundancies. For the alternative versions of Figures 5–8 showing the relation between sensitivity and differences in instantaneous-frequency speed along the cochlear tonotopic map, we refer the interested reader to the Supporting Information.

3 | 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 cues (sweep-direction experiment) and those providing both
acoustic ENV and acoustic TFS cues (sweep-velocity experiment). In the sweep-direction experiment, discrimination was between SCHR− and SCHR+ complexes with C values only differing in sign, whereas in the sweep-velocity experiment, SCHR complexes with different C values, which could have different duty cycles, were discriminated. In both experiments neural responses were evaluated by two different vR analyses, using a time constant of 1 ms or a time constant of 1/f0, as well as an analysis based on mean rate.

3.1 Neural responses to sweeps that differed in direction

Representative examples of responses from two AN fibers and four IC neurons are shown in Figures 3 and 4, respectively. The AN fiber in Figure 3a had BF = 400 Hz, whereas the example in Figure 3b had BF = 1800 Hz. Both fibers phase-locked to f0, as evident in the raster plots, for f0 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 Figure 3a was higher at low f0 values (50 and 100 Hz) than at higher f0 values, when based on the vR analysis with τ = 1 ms (Figure 3a, yellow). The vR analysis with τ = 1/f0 (green) and rate analysis (black) yielded low d’ values for all SCHR complexes. For the example fiber in Figure 3b, the d’ based on vR analyses with τ = 1 ms was similar to the other example fiber; however, in some cases, d’ based on vR analyses with τ = 1 ms slightly exceeded 1 at f0 = 200 Hz and f0 = 400 Hz. The d’ values based on vR analyses with τ = 1/f0 and rate representations exceeded 1 for selected conditions, but were mostly below 1 again. The relative values of d’ obtained with the vR analyses with τ = 1 ms, vR analyses with τ = 1/f0, and rate analyses varied between AN fibers.

IC neurons had diverse selectivity for sweep direction; the responses in Figure 4 are for three different MTF types. Figure 4a shows an example of a neuron with a band-pass MTF that responded vigorously to both SCHR+ and SCHR− complexes across all conditions, with overall rates decreasing at higher f0 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 although the d’ values for the vR analyses with τ = 1 ms decreased to 0 for f0 of 200 and 400 Hz, neural discrimination was still possible based on vR analyses with τ = 1/f0 up to 200 Hz.

Figure 4b shows a band-pass neuron with striking selectivity for the SCHR complexes that varies with f0. It is clear that the responses of this neuron would allow discrimination between the SCHR+ and SCHR− complexes

![Figure 3](image-url)  
**Figure 3** Example responses of two different AN fibers (a,b) to positive (SCHR+, red) and negative (SCHR−, blue) SCHR complexes with different f0 (50, 100, 200, and 400 Hz) and C values (±0.25, ±0.5, ±0.75, ±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 vR spike distance with a time constant τ = 1 ms in yellow, and d’ based on the vR spike distance with a time constant of 1/f0 in green. Dot-dashed lines indicate d’ values between 0 and 1. Note that d’ above 1 indicates above-threshold sensitivity. Within each f0 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 f0 and with decreasing C value for each f0 (see also Figure 1).
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).

Figure 4c 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, Figure 4d illustrates the responses of a neuron with a hybrid MTF, that is, 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 to 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 tested conditions.

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

Responses of 22 AN fibers from two male and three female gerbils were analyzed to estimate discrimination performance that could be achieved on the basis of each AN fiber’s responses. BFs ranged from 400 to 4600 Hz. Note that this BF range reflects a deliberate bias toward the range where AN fibers are known to represent stimulus TFS via phase locking (Versteegh et al., 2011).

For spike trains analyzed with $\tau = 1$ ms (Figure 5a), 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, that is, 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. For the spike trains analyzed with $\tau = 1/f_0$ (Figure 5c), the mean $d'$ values for nearly all conditions were lower than for the results with $\tau = 1$ ms and almost uniformly below 1. Finally, Figure 5e 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 identified $f_0$ as the only significant factor that influenced the neural discrimination sensitivity for both $vR$ analyses, with a tendency toward higher $d'$ values for lower $f_0$s. The rate-based analysis showed no significant effects of neither $f_0$ nor $C$ value. Additional GLMM ANOVAs, which included the fibers’ BF class, revealed that for all discrimination strategies, BF class did not affect the sensitivity.

In summary, the spiking patterns of AN fibers, analyzed at high temporal resolution using $\tau = 1$ ms, were best suited to discriminate the sweep direction of SCHR complexes for the lower $f_0$s of 50 and 100 Hz (Figure 5a,c,e).
and broke down at $f_0$ of 400 Hz. In contrast, both rate-based coding and coding based on $\tau = 1/f_0$ provided poor discrimination of sweep direction at all $f_0$s and $C$ values tested. Thus, when the only cue was sweep direction, changes in spike timing analyzed using $\tau = 1$ ms could explain discrimination, especially for low $f_0$s. Changes in rate and in spike timing based on $\tau = 1/f_0$ for most comparisons could not explain discrimination.

### 3.3 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 vR-based analyses were limited to conditions for which the mean response rate was at least 1 spike/s 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.

For spike trains analyzed with $\tau = 1$ ms (Figure 5b), 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 vR analyses with $\tau = 1$ ms by IC neurons. Note, however, that some individual neurons were highly sensitive to sweep direction (e.g., Figure 4a). When the spike trains were analyzed with $\tau = 1/f_0$ (Figure 5d), $d'$ reached slightly higher values than for the analysis with $\tau = 1$ ms, especially for $f_0 = 50$ and 100 Hz, with absolute $C$ values of 0.75 and 1. In contrast to both vR-based analyses, the rate-based sensitivity values (Figure 5f) were generally above the threshold of $d' = 1$, with the maximum mean $d' = 2.11$ for $f_0 = 100$ 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 used in the analyses. The impact of $f_0$, however, was highly significant. Subsequent GLMM ANOVAs, carried out separately for the two $vR$-based and the rate-based neural discriminations, 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 the $vR$ analysis with $\tau = 1$ ms, the BMF class had a significant influence on sensitivity ($p = 0.017$), with neurons having a BMF up to 75 Hz exhibiting the lowest $d'$ values. For the $vR$ analysis with $\tau = 1/f_0$, the MTF type had a significant influence on $d'$ ($p = 0.002$). The pairwise post hoc comparisons revealed a maximum sensitivity for the band-reject MTF-type neurons. For rate-based analysis, both BMF class and MTF type influenced the sensitivity ($p = 0.007$ and $p < 0.001$, respectively). Sensitivity generally increased with increasing BMF class, and a pairwise post hoc comparison showed that the band-reject MTF type had the highest sensitivity. Tables with the full ANOVA results including all interactions are presented in the Supporting Information.

In summary, the sensitivity of sweep-direction discrimination by rate coding was typically above threshold ($d' > 1$) in the IC, whereas the average $vR$-based sensitivity was mostly below threshold. When the only acoustic cue was sweep direction, the $d'$ values were highest for rate and spike trains analyzed with $\tau = 1/f_0$, whereas the $\tau = 1$ ms-based $d'$ values were rarely higher than 1. Comparing the two stages of neural processing investigated here, the mean $d'$ values of IC neurons were highest when based on rate representations, whereas AN fibers achieved the highest mean $d'$ values for $vR$ analyses with $\tau = 1$ ms.
3.4 Behavioral discrimination of sweep direction by gerbils and humans

The gerbils’ sensitivity for behaviorally discriminating stimuli that differed in sweep direction, as a function of the absolute C value, is shown in Figure 5g. 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 $d’ = 0.64$ for $f_0 = 400 \text{ Hz}$ and $|C| = 0.25$ to mean $d’ = 2.67$ for $f_0 = 50 \text{ Hz}$ and $|C| = 0.75$. The GLMM ANOVA showed significant main effects of $C$ value and $f_0$. Thus, behavioral discrimination depended on $C$ value, whereas neural discrimination did not.

In order to provide a limited comparison between the behavioral results in gerbils and related studies in human listeners (e.g., Drennan et al., 2008; Lauer et al., 2009), the sweep-direction experiment was carried out in four
human subjects using 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 $\mid C \mid = 0.25$ to mean $d' = 4.14$ for $f_0 = 50$ Hz and $\mid C \mid = 0.50$ (Figure 5b). A GLMM ANOVA revealed a significant main effect of $C$ value 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 ($p < 0.001$). These results were generally consistent with previously reported results in human listeners (e.g., Drennan et al., 2008; Lauer et al., 2009). Note that the human testing was only carried out for the sweep-direction experiment, which had fewer test conditions than the sweep-velocity experiment and for which comparison data were available (Dooling et al., 2002).

3.5 | 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 $C$ values 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 $C$ values. 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 $C = 1$ (Figure 6a,d,f). The mean $d'$ values based on vR analyses with $\tau = 1$ ms exceeded 1 for stimuli with $f_0 = 50$ Hz and negative $C$ values (Figure 6a). For all other $f_0$s, mean $d'$ values were below 1 for most stimuli, indicating that spike trains analyzed with $\tau = 1$ ms in the responses to reference and target stimuli did not support neural discrimination. For spike trains analyzed with $\tau = 1/f_0$ (Figure 6d), $d'$ values for target stimuli with $f_0$s of 50 and 100 Hz and absolute $C$ values of 0.25 and 0.5 were higher compared to the $\tau = 1$ ms-based vR analyses. The further the absolute $C$ value deviated from 1, the better was the neural discrimination of the AN fibers, indicating that this period-related metric strongly reflects differences in duty cycle. The rate-based analysis (Figure 6g), similar to the vR analysis with $\tau = 1/f_0$, showed best neural discrimination for low $f_0$s (i.e., $f_0 = 50$ Hz and 100 Hz) and $C = -0.25$ and 0.25 (Figure 6f). As expected, mean $d'$ values decreased for smaller differences in absolute value of $C$ between reference and target, that is, stimuli with more similar duty cycles were harder to discriminate. A GLMM ANOVA revealed that $f_0$ and $C$ value both had a highly significant effect on the neural discrimination of sweep velocity, based on $\tau = 1/f_0$ and rate. Furthermore, a significant effect of only $f_0$ on $d'$ values based on vR analyses with $\tau = 1$ ms was confirmed.

Results for the AN fiber responses analyzed using a reference stimulus with $C = -1$ (Figure 7a,d,f) were similar to those described above for $C = 1$. Here, vR analysis with $\tau = 1$ ms (Figure 7a) resulted in mean $d'$ values exceeding 1 for all comparisons of SCHR complex stimuli with $f_0 = 50$ Hz. For $f_0 = 200$ and 400 Hz, most $d'$ values were below 1, indicating that these discriminations were not possible, based on spike trains analyzed with $\tau = 1$ ms. Results for $f_0 = 100$ Hz were intermediate. Spike trains analyzed with $\tau = 1/f_0$ (Figure 7d) showed that $d'$ values were again higher compared with those analyzed with $\tau = 1$ ms and somewhat lower than those analyzed based on rate. Again, the highest sensitivity was found for neural discriminations in which target and reference differed most in duty cycle. The rate-based analysis (Figure 7f) resulted in higher mean $d'$ values than either of the vR-based analyses, for most comparisons. 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 $d'$ values were maximal when the difference between absolute values of $C$ was largest, that is, for target $C$ values of 0.25 and $-0.25$, as was also apparent for spike trains analyzed with $\tau = 1/f_0$. This again indicates that the sensitivity for discrimination was strongly related to the difference in duty cycle between reference and target. Consistent with these qualitative observations and similar to the results obtained with a reference of $C = 1$, a GLMM ANOVA showed that discrimination based on mean rate as well as for analyses based on $\tau = 1/f_0$ was significantly affected by both $f_0$ and $C$ values. For $d'$ values obtained with vR analyses with $\tau = 1$ ms, only $f_0$ had a significant effect. An additional GLMM ANOVA including $f_0$, $C$ values, and BF class for the rate-based analyses and vR-based analyses with $\tau = 1/f_0$ revealed the BF class as a further significant factor affecting sensitivity (both $p < 0.001$; see Supporting Information), such that fibers with BF > 1.850 kHz achieved higher sensitivities, whereas the significant effect of $f_0$ and $C$ values remained.

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 based on vR analysis with $\tau = 1/f_0$ and mean rate, whereas the responses based on vR analysis with $\tau = 1$ ms were less discriminable. Furthermore, presenting opposing sweep directions together with a velocity difference resulted in spike trains that were slightly
more discriminable based on vR analysis with \( \tau = 1 \) ms than were stimuli that had the velocity cue alone. This difference is reflected in an asymmetry of the \( d' \) distribution and can be seen in Figure 6a, for \( f_0 = 50 \) Hz and reference \( C = 1 \), where \( d' \) values were higher for the discrimination with negative \( C \) values, whereas 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 was primarily represented by changes in spike trains when analyzed with high temporal resolution (\( \tau = 1 \) ms), whereas sweep velocity and duty cycle were represented by changes in spike trains revealed by analyses with lower temporal resolution (\( \tau = 1/f_0 \) and rate).

3.6 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 vR-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. Sensitivity based on both vR analyses and rate are presented separately for reference stimuli with \( C = 1, -1, \) and 0.

For the reference \( C = 1 \) (Figure 6), the mean discrimination sensitivity increased with longer integration times. For spike trains analyzed with \( \tau = 1 \) ms, the best mean sensitivity was \( d' = 1.00 \) for \( f_0 = 50 \) Hz (Figure 6b). For spike trains analyzed with \( \tau = 1/f_0 \) (Figure 6e), for \( f_0 \) of 50 and 100 Hz, the \( d' \) values were typically highest for the target \( C \) values with the opposite sign to the reference. The \( d' \) values were mostly below 1 for \( f_0 = 200 \) and 400 Hz. Rate-based discrimination resulted in \( d' > 1 \) for every \( f_0 \), with the highest mean \( d' = 2.37 \) for \( f_0 = 100 \) Hz (Figure 6g). Thus, in the IC, discrimination of SCHR complexes with different \( C \) values from the reference \( C = 1 \) was generally possible based on spike trains analyzed with \( \tau = 1/f_0 \) and rate representation, but not based on spike trains analyzed with \( \tau = 1 \) ms. A GLMM ANOVA showed that \( f_0 \) affected \( d' \) values for every discrimination analysis. The \( C \) value, however, only had a significant impact on \( d' \) values based on spike trains analyzed with \( \tau = 1/f_0 \) and rate-based discrimination, which also were the only metrics reaching \( d' \) values >1. Additional GLMM ANOVAs were carried out for neural discrimination based on either vR analyses or rate-based analysis, including BF class, BMF class, and MTF type as dependent variables (see Supporting Information). The results were consistent with the main effects for \( f_0 \) and \( C \) values shown above each of the panels in Figures 6–8.

The BF class affected discrimination based on spike trains analyzed with \( \tau = 1/f_0 \) (\( p = 0.011 \)) and rate-based discrimination (\( p = 0.012 \)), with higher \( d' \) values for neurons with a BF > 1.850 kHz. The modulation-based parameter BMF class influenced sensitivity for all metrics ([\( p < 0.001 \]) for \( \tau = 1 \) ms, [\( p < 0.001 \]) for \( \tau = 1/f_0 \), and [\( p = 0.005 \]) for rate). For spike trains analyzed with \( \tau = 1 \) ms, neurons with a BMF up to 75 Hz reached lowest \( d' \) values, whereas a generally increasing sensitivity with increasing BMF class was observed for spike trains analyzed with \( \tau = 1/f_0 \) and rate-based discrimination. The MTF type had an effect on spike trains analyzed with \( \tau = 1 \) ms (\( p < 0.001 \)), indicating that band-pass MTF-type neurons could discriminate best between SCHR complexes of different sweep velocity, whereas hybrid MTF-type neurons showed lowest \( d' \) values. Furthermore, the MTF type also affected the sensitivity in the rate-based discrimination (\( p < 0.001 \)), such that neurons with band-reject MTF types achieved the largest \( d' \) values.

For the reference \( C = -1 \), a similar pattern as for the reference \( C = 1 \) was observed (Figure 7b.e.g). The sensitivity also increased with longer integration time and for spike trains analyzed with \( \tau = 1 \) ms, no \( d' > 1 \) were achieved. Thus, also for \( C = -1 \), discrimination related to high temporal resolution (\( \tau = 1 \) ms) of SCHR complexes was not possible (Figure 7b). Discrimination related to lower temporal resolution (\( \tau = 1/f_0 \) and rate) resulted in sensitivities above threshold (Figure 7e.g). The trend of lower sensitivity with lower differences between absolute values of reference and target \( C \) can also be seen for these neural discriminations. The GLMM ANOVAs revealed an effect of \( f_0 \) for every discrimination and an effect of \( C \) value for spike trains analyzed with \( \tau = 1/f_0 \) as well as rate-based discrimination. The additional GLMM ANOVAs showed a BF-class dependency only for discrimination based on vR analyses ([\( p < 0.001 \]) for \( \tau = 1 \) ms and [\( p = 0.001 \]) for \( \tau = 1/f_0 \), with neurons having a BF above 1.850 kHz showing higher sensitivity. The BMF class affected the sensitivity for spike trains analyzed with \( \tau = 1 \) ms (\( p < 0.001 \)) and rate-based discrimination (\( p < 0.001 \)), reflecting that for vR analysis with \( \tau = 1 \) ms, neurons with a BMF up to 75 Hz discriminated worst, as was apparent for neural discrimination with the reference of \( C = 1 \). For rate-based discrimination, neurons showed increasing \( d' \) values with increasing BMF. The MTF type affected discrimination based on vR analysis with \( \tau = 1 \) ms (\( p < 0.001 \)), discrimination related to lower temporal resolution (\( \tau = 1/f_0 \)) (\( p < 0.001 \)), and rate-based coding (\( p < 0.001 \)), with band-pass MTF types showing the highest sensitivity for coding with a high temporal resolution (\( \tau = 1 \) ms) and
band-reject MTF type for coding with a lower temporal resolution (τ = 1/f0) and rate coding.

Finally, for the reference C = 0, sensitivity also increased with longer integration time (Figure 8a–c). Sensitivities were mostly below 1 for spike trains analyzed with τ = 1 ms and the sensitivities for spike trains analyzed with τ = 1/f0 and rate decreased with decreasing differences in C value. The GLMM ANOVAs showed a significant effect of f0 for all discrimination types and an effect of C value for spike trains analyzed with τ = 1/f0 and rate. The additional GLMM ANOVAs revealed a main effect for BF class for spike trains analyzed with τ = 1/f0 (p = 0.001) and for rate-based discrimination (p = 0.013), with neurons having larger BMFs achieving larger d’ values. The modulation-related parameter BMF class influenced the sensitivity based on vR analysis with τ = 1 ms (p < 0.001) and rate (p = 0.014). Neurons having BMFs up to 75 Hz obtaining lowest d’ values for the temporal discrimination, whereas neurons with BMFs of 150 Hz and above achieved largest d’ values for rate-based analysis. The MTF type affected temporal representation when analyzed with τ = 1 ms (p < 0.001) and rate-based representation (p = 0.005), where the band-pass MTF-type neurons reached highest sensitivity values for coding with a high temporal resolution (τ = 1 ms), whereas for rate-based coding, neurons with a band-reject MTF achieved highest d’ values.

In summary, when velocity was added as a cue to SCHR stimuli of opposing sweep direction, spike patterns of IC neurons analyzed with τ = 1 ms revealed little temporally distinct information, but supported the discrimination primarily by their mean rate, and to a lesser extent by information at lower temporal resolution (τ = 1/f0). Combining both types of cues, sweep velocity and sweep direction, resulted in higher sensitivities in responses of IC neurons as compared with d’ values in response to stimuli that differed only in sweep direction. However, in both experiments, rate-based analyses and vR analyses based on τ = 1/f0 revealed greater sensitivity than analysis based on τ = 1 ms. In contrast, the ranking of neural metrics of AN responses did change when the new cues were introduced in the sweep-velocity experiment.

3.7 | Discrimination of sweep velocity based on behavioral responses by gerbils

Figures 6c, 7c, and 8d show the behavioral sensitivity, d’ ([mean, ±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 ≤ 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 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], 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, that is, a difference in duty cycle. The discrimination sensitivity between SCHR+ and SCHR− with C = [1] for f0 < = 100 Hz were lower in the sweep-velocity experiment than in the sweep-direction experiment (compare Figure 5g with Figures 6c and 7c). A possible explanation for this difference is that the training in the sweep-velocity experiment caused the gerbils to focus mainly on ENV cues in the stimuli, whereas the training in the sweep-direction experiment caused the gerbils to focus mainly on TFS cues in the stimuli.

3.8 | Correlations between behavioral and neural discrimination

Correlating the sensitivity derived from behavioral responses and the sensitivity derived from neuronal responses using vR- or rate-based analyses reveals their relations (Figures 9 and 10). A high correlation coefficient indicates a linear relationship between the behavioral and neural response measures, unless the slope is not significantly different from zero, which would indicate no relation. If the behavioral and neural sensitivities match quantitatively, the data points will cluster around the diagonal. Data points below the diagonal indicate that neural response metrics are more sensitive than behavior. Data points above the diagonal indicate that behavior is more sensitive than the neural response metrics. These indicators were used to examine which of the neural response patterns in the AN or the IC better explained behavioral performance.

3.9 | Sweep-direction experiment

We computed the correlation between the behaviorally obtained d’ values and average d’p values obtained
when pooling (i.e., randomly combining) the sensitivity of pairs of AN fibers or IC neurons, respectively, for all combinations of \( C \) values and \( f0 \). This analysis was done separately for vR-based analyses with time constants of either 1 ms or \( 1/f0 \) and rate-based analyses (Figure 9). For AN fibers and IC neurons, both vR-based \( d'_{p} \) values were highly correlated with the behaviorally obtained sensitivity (Figure 9a-d), whereas rate-based \( d'_{p} \) was less strongly correlated with behavior than were the vR-based \( d'_{p} \) values (Figure 9e,f). There was a good match between the sensitivity of AN fiber responses analyzed with a 1-ms time constant and behavior, as indicated by the data points clustering closely around the diagonal line, and a high correlation coefficient (Figure 9a). In the other comparisons involving AN responses, the neural sensitivity tended to be lower than the behavioral sensitivity (Figure 9c,e), and the correlation coefficient was higher for the vR-based discrimination than for the rate-based discrimination. The comparisons involving IC responses using temporal vR-based response measures also showed a lower neural than behavioral sensitivity, whereas the correlation was still high (Figure 9b,d). In the same comparisons, it was also obvious that IC neurons can hardly discriminate the sweep direction based on temporal representations for stimuli with \( f0 \)s of 200 or 400 Hz, whereas the gerbils’ behavior showed discrimination for SCHR complexes with \( f0 \)s up to 200 Hz. As judged by the value of the correlation coefficient, the worst match was observed between \( d'_{p} \) of the rate response of IC neurons and behavior. However, the sensitivity based on the rate responses of the IC neurons was considerably higher than behavioral sensitivity, that is, the data points were generally below the diagonal (Figure 9f). This result indicated that IC neurons have a higher average sensitivity than AN fibers in discriminating the direction of SCHR complexes based on the rate response, but this high sensitivity was not utilized by the circuitry reflected in the behavioral discrimination. This discrepancy between neural and behavioral sensitivity was evidenced, for example, in the results for \( f0 = 400 \) Hz, a condition for which the IC neurons clearly outperform the gerbil’s behavior. Thus, based on the correlation analysis, differences in rate appear less suitable to explain the behavioral discrimination of the sweep direction of SCHR complexes, although the sensitivity of the IC neurons would be sufficiently high to account for the behavioral performance.

To summarize, the correlations between neural \( d' \) values and behavioral sensitivity were strongest for both temporal metrics, for both AN and IC, whereas the rate representation was less strongly correlated with behavior. This result suggests that temporal response measures reflect sweep direction better than rate-based response measures for a task in which discrimination relies mainly on stimulus TFS cues.

**FIGURE 9** Comparison of mean behavioral discrimination sensitivity, \( d' \), with neural pooled sensitivity, \( d'_{p} \), obtained in the sweep-direction experiment. Correlations for AN fibers (\( N = 16 \) conditions; a,c,e) and neurons in the IC core (\( N = 16 \) conditions; b,d,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 \( d' \); a,b), a time constant \( \tau \) of \( 1/f0 \) (ENV-related \( d' \); c,d) and \( d'_{p} \) based on mean discharge rate (e,f). Data for different \( f0 \)s 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.
3.10 | 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 (Figure 10). The correlation between the sensitivity of AN fibers obtained with a time constant of 1 ms in the $vR$ analysis and the behavioral sensitivity was considerably lower than the correlations observed with the sensitivity obtained with the $vR$ analysis with a time constant of $1/f_0$ or the response rate. This result is in contrast to the result for the sweep-direction experiment, in which the $vR$ analysis with $\tau = 1$ ms provided the best match to the behavioral sensitivity. When the duty cycle differs between stimuli, as is the case in the sweep-velocity experiment, response measures relying on longer integration times (i.e., $1/f_0$ or the total stimulus in the rate response) correlate better with the behavioral response measure. In addition, the data points in the correlations involving AN fiber response measures that were based on the temporal $vR$ analyses were generally above the diagonal, whereas data points for the rate-based AN response measure were slightly closer to the diagonal (Figure 10k,l). These results suggest that, for discrimination of stimulus ENV cues in the sweep-velocity experiment, the most sensitive AN response metrics were averaged rate and the temporal metric based on $1/f_0$. 

**FIGURE 10** 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; a,b,f,g,k,l) and IC ($N = 32$ conditions; c,d,e,h,i,j,m,n,o). Furthermore, correlations of the same behavioral data are shown with three different neural measures: temporally based $d'_{p}$ and of $vR$ spike distance with a time constant $\tau$ of 1 ms (TFS-related $d'$; a–e), a time constant $\tau$ of $1/f_0$ depending on the fundamental period (ENV-related $d'$; f–j) and $d'_{p}$ based on mean discharge rate (k–o). Panels with the same reference $C$ value are grouped within one column for AN ($C = 1$: a,f,k; $C = -1$: b,g,l) and IC ($C = 1$: c,h,m; $C = -1$: d,i,n; $C = 0$: e,j,o). Data for different $f_0$s 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.
The correlations between behavioral sensitivity and sensitivity of IC neurons resemble in some aspects the observations made for AN fibers. Sensitivities derived from the $vR$ analyses were generally lower than the behavioral sensitivity, as indicated by the majority of the data points lying above the diagonal line (Figure 10c,d,e,h,i,j). Nevertheless, the sensitivity of IC neurons derived from the $vR$ analyses and the sensitivity measured in behavior were significantly correlated, probably because IC neurons generally had a low sensitivity to discriminate stimuli with $f_0$s of 200 and 400 Hz (as indicated by clustering of those data points next to the y-axis; Figure 10c,d,e,h,i,j). Sensitivities of IC neurons derived from the rate response, however, were mostly superior to behavioral sensitivity (i.e., lie below the diagonal), and the values of the sensitivity of IC neurons were generally larger than those of the AN fibers (Figure 10m–o). Such a difference may result from convergent input to the IC neurons and integration across different frequency channels in the auditory pathway. Consistent with the increased information that would accompany such an integration, it must be noted that neural $d’$ values derived from IC neurons’ rate responses often were capped at a maximum value.

To summarize, when acoustic TFS was the primary cue for discrimination (sweep-direction experiment), correlations between neural and behavioral $d’$ values were strongest for the two temporal metrics, for both AN and IC. When both TFS and ENV cues were available for discrimination, as well as differences in duty cycle (sweep-velocity experiment), correlations between behavioral and neural $d’$ values were strongest for AN metrics based on average rate or on the temporal analysis with lower resolution ($\tau = 1/\bar{f}$) and for IC metrics based on temporal analyses with both low and high ($\tau = 1 \text{ ms}$) resolutions. However, for IC neurons, $d’$ values based on average rates were higher than behavioral $d’$ values for both sweep-direction and sweep-velocity experiments.

### 4 Discussion

This study explored the potential contributions of neural temporal and rate-based coding of the ENV and TFS cues in acoustic stimuli to explain behavioral discrimination of SCHR complexes. It is useful to distinguish separate questions when comparing the $d’$ values 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?

#### 4.1 Mechanisms underlying neural discrimination

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

SCHR complexes that differ only in the sign of $C$ have similar ENVs and therefore would be expected to be discriminated primarily based on TFS cues. However, cochlear filtering modifies the representation of the ENV due to dispersive properties of the BM traveling wave and cochlear compression (Smith et al., 1986). These factors predict different ENVs of BM responses, such that the down-sweeping SCHR+ complex would elicit a more modulated, peakier response than SCHR– (Carloyyn & Datta, 1997), as experimentally confirmed for high-frequency BM locations (Recio & 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 shapes of the ENV on the BM and RMS amplitudes of BM responses to the two SCHR polarities (Recio & Rhode, 2000) would be expected to elicit different mean discharge rates in the sweep-direction experiment. However, sensitive discrimination based on average AN rates was not evident (Figure 5e), plausibly explained by the low-frequency bias of our AN sample. Consistent with that observation, diminished differences between masked thresholds for tones in SCHR complexes with opposite polarities were observed at 500 Hz and below in human subjects (Oxenham & Dau, 2001). 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 apical BM response ENVs 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, small differences in mean discharge rates evoked by the two SCHR polarities were found in the AN and ventral cochlear nucleus only at CFs higher than 3–4 kHz (Recio, 2001).
Our analysis based on pooling the sensitivity of two randomly selected neurons revealed that temporal responses, rather than rate responses, of AN fibers predicted the gerbils’ behavioral discrimination in the sweep-direction experiment (Figure 9). We conclude that despite known differences in BM response ENVs that arise from cochlear filtering, discrimination of the sweep direction of SCHR complexes with opposite-signed C values was largely carried by the representation of the stimulus TFS in AN fibers with low BF. Only when additional cues were introduced—velocity and duty cycle in the sweep-velocity experiment—were the AN rate and the representation of spikes analyzed with lower temporal resolution ($\tau = 1/f_0$) potentially more informative for the gerbils’ behavioral decision (Figure 10). Additional information about the sweep direction may be derived by patterns of AN responses across fibers tuned to different BFs. In response to 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. Our additional analysis relating sensitivity to a measure of differences in the speed and direction of the change in instantaneous frequency along the frequency map of the cochlea (see Supporting Information) supports the view that excitation pattern shifts can also be used for discriminating SCHR stimuli, yielding similar results to the analysis based on C value and $f_0$.

Neurons at higher levels of the auditory pathway, as a rule, have weaker phase locking to acoustic TFS; however, phase locking in the IC to the stimulus 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 (Kim et al., 2020; Schreiner & Langner, 1988; reviewed in 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 & 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 & 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, Figure 4 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. In addition, IC neurons likely receive inputs originating from AN fibers with a range of BFs. Thus, by combining information from many inputs, IC neurons can be more sensitive to sweep direction and velocity than AN fibers, as evidenced by their high sensitivity compared with AN fibers, based on their rate response. Pooling the rate-based sensitivity of randomly selected pairs of IC neurons allowed better discrimination of SCHR stimuli than that observed in the behavior (Figures 9 and 10).

In the analysis of the sensitivity of both AN fibers and IC neurons, we observed significant effects of BF (ANOVA of AN and IC responses including the factor BF class) and of BMF and of the type of MTF (ANOVA of IC responses including the factors BMF class and MTF type). In all cases showing significant effects of BF, both in AN and IC, higher BFs (fibers and neurons with BF > 1850 Hz) were associated with higher sensitivity. Interestingly, the SCHR stimuli elicited a stronger response (i.e., more spikes) in high-BF neurons than in low-BF neurons. This relation was also found for the other factors, BMF class and MTF type. If there was a significant difference in the pairwise comparisons of the sensitivity of different MTF types or BMF classes, it associated with significant differences in the response strength, that is, neurons that responded more vigorously showed a higher sensitivity for discrimination. This common observation could be explained by a very general effect: If a comparison involves responses with more action potentials for at least one condition, then the representation of the stimuli will be more distinct from a background of neural noise, and more action potentials may provide a better sample for statistical comparisons.

### 4.2 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 sensitivities (see Figures 9 and 10 for sensitivity achieved by pooling the sensitivity of two randomly selected neurons). This observation is not unique; it has been reported previously that the sensitivity of individual neurons can be much higher than the behavioral
sensitivity for the same discrimination (e.g., Carney et al., 2014; Johnson et al., 2012; Klink & Klump, 2004). 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 neural $d'$ and behavioral sensitivity support this prediction. If the average neural sensitivity was lower than the behavioral sensitivity, then the sensitivity of a single neuron would not be sufficient, and the information provided by a number of neurons must be combined to explain behavioral sensitivity. Our results indicate that pooling the performance of a small number of neurons (e.g., two AN fibers in the sweep-direction experiment; Figure 9a) 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 behavioral performance, not precluding that additional brain areas (e.g., cortex) are involved.

4.3 | 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 could not explain the birds’ discrimination of sweep direction. The same was true for gerbils, when comparing CAP measurements reported by Dooling et al. (2002) with the behavioral performance observed in the present study. Thus, CAP peak amplitude is not a valid metric in this context, and we suggest this is due to the CAP being dominated by the highly synchronized responses of AN fibers (Bourien et al., 2014; Özdamar & Dallos, 1978). For SCHR stimuli, the temporal synchronization of activity across cochlear 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 with SCHR+ (that actually elicit peekier responses at high-frequency BM locations).

Gerbils are often used as an animal model for human peripheral processing due to low-frequency hearing that is similar 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 inner hair cells (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 with humans (35 mm, Bredberg, 1968; 3500 inner hair cells, Wright et al., 1987; 0.01–16 kHz; ISO 389-7, 2019). 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 (Klinge et al., 2010; Shera et al., 2010) and differences of neural tuning in the ascending auditory pathway.

5 | CONCLUSIONS

The present study investigated whether the information conveyed by single AN fibers could predict behavioral discrimination of SCHR stimuli and whether integration of the information at the midbrain level resulted in better discrimination by single IC neurons. Furthermore, this study explored which representation of SCHR stimuli in the neural responses provided for a better match with behavior. In general, in both AN fibers and IC neurons, discrimination of SCHR stimuli was poor for $f_0$ of 400 Hz, and sensitivity increased with decreasing $f_0$. In the sweep-direction experiment, discrimination based on AN fiber temporal responses analyzed with a 1-ms time constant was highly correlated with the gerbils’ perception, and the information of two AN fibers combined was sufficient to explain behavioral discrimination. IC neurons, however, showed the most sensitive discrimination of SCHR stimuli based on their rate and were superior compared with AN fibers, suggesting an enhancement of sensitivity that could be achieved by combining multiple inputs from the periphery and integrating over a longer time period. Although the sensitivity of IC neurons for discriminating sweep direction based on rate was less well correlated with perception than discrimination based on temporal response measures, the average rate-based sensitivity based on pairs of neurons was more than sufficient to explain behavioral performance. In the sweep-velocity experiment, IC neurons also represented differences between SCHR stimuli best based on rate, showing a higher average sensitivity than was observed in the gerbil behavior. Also, in AN fibers, the average rate-based sensitivity matched the perception better than sensitivity based on temporal analysis. Thus, in AN responses, the stimulus TFS was represented by temporal spike patterns, whereas stimulus ENV was represented.
by rate. In the IC, stimulus TFS and ENV were both represented best by rate.

ACKNOWLEDGEMENTS
This work was supported by the Deutsche Forschungsgemeinschaft (DFG) priority program “PP 1608,” the DFG Cluster of Excellence EXC 1077/1 “Hearing4all” (Project ID 390895286), National Institutes of Health (NIH-R01-001641) (LHC), and the Hanse-Wissenschaftskolleg. We thank Susanne Groß for assistance with the behavioral experiments, Lichun Zhang and Sandra Tolnai for assistance with the electrophysiological recordings, and Elisabeth Groscurth and David Cameron for assistance with histology and reconstructions.

CONFLICT OF INTEREST
The authors declare no competing financial interests.

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.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15744.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are openly available in Dryad at https://doi.org/10.5061/dryad.0gb5mkm17.

ORCID
Rainer Beutelmann https://orcid.org/0000-0001-9708-4113
Laurel H. Carney https://orcid.org/0000-0002-4729-5702
Christine Koeppl https://orcid.org/0000-0003-0729-709X
Georg M. Klump https://orcid.org/0000-0003-2501-8090

REFERENCES
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 (Suppl 2), 1131–1139. https://doi.org/10.1121/1.1912474
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, 87–101. https://doi.org/10.1016/0042-6989(71)90033-2

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, EL71–EL77. https://doi.org/10.1121/1.4903921
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, 1025–1039. https://doi.org/10.1152/jn.00738.2013
Bredberg, G. (1968). Cochlear structure and hearing in man. In A. V. S. De Reuck & J. Knight (Eds.), *Ciba Foundation symposium - hearing mechanisms in vertebrates* (pp. 126–142). John Wiley & Sons, Ltd.
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, 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. *Journal of the Association for Research in Otologyngology*, 18, 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. *Journal of the Association for Research in Otologyngology*, 19, 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, 2384–2391. 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. *The Journal of Neuroscience*, 34, 1306–1313. https://doi.org/10.1523/JNEUROSCI.3031-13.2014
Cedolin, L., & Delgutte, B. (2010). Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve. *The Journal of Neuroscience*, 30, 12712–12724. https://doi.org/10.1523/JNEUROSCI.6365-09.2010
Cheal, M. (1986). The gerbil: A unique model for research on aging. *Experimental Aging Research*, 12, 3–21. https://doi.org/10.1080/03610738608259430
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, 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 deccerebrate gerbil. *Journal of Neurophysiology*, 75, 1411–1431. https://doi.org/10.1152/jn.1996.75.4.1411
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, 748–759. https://doi.org/10.1121/1.1494447
Drennan, W. R., Longnion, J. K., Ruffin, C., & Rubenstein, J. T. (2008). Discrimination of Schroeder-phase harmonic
complexes by normal-hearing and cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology*, 9, 138–149. https://doi.org/10.1007/s10162-007-0107-6

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, 2327–2341. https://doi.org/10.1152/jn.00326.2006

Escañilla, 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. *The Journal of Neuroscience*, 35, 6284–6297. https://doi.org/10.1523/JNEUROSCI.3526-14.2015

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Heil, P., Neubauer, H., Irvine, D. R. F., & Brown, M. (2007). Spontaneous activity of auditory-nerve fibers: Insights into stochastic processes at ribbon synapses. *The Journal of Neuroscience*, 27, 8457–8474. https://doi.org/10.1523/JNEUROSCI.1512-07.2007

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

ISO 389:7-2019. (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.

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America*, 68, 1115–1122. 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, 3325–3341. https://doi.org/10.1152/jn.00812.2011

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews*, 84, 541–577. https://doi.org/10.1152/physrev.00029.2003

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, 1037–1051. https://doi.org/10.1152/jn.1994.71.3.1037

Joris, P. X., & Yin, T. C. T. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America*, 91, 215–232. https://doi.org/10.1121/1.402757

Keine, C., & Ruesbsamen, R. (2015). Inhibition shapes acoustic responsiveness in spherical bushy cells. *The Journal of Neuroscience*, 35, 8579–8592. https://doi.org/10.1523/JNEUROSCI.0133-15.2015

Kim, D. O., Carney, L. H., & Kuwada, S. (2020). Amplitude modulation transfer functions reveal opposing populations within both the inferior colliculus and medial geniculate body. *Journal of Neurophysiology*, 124, 1198–1215. https://doi.org/10.1152/jn.00279.2020

Kittel, M., Wagner, E., & Klump, G. M. (2002). An estimate of the auditory-filter bandwidth in the Mongolian gerbil. *Hearing Research*, 164, 69–76. https://doi.org/10.1016/S0378-5955(01)00411-7

Klinge, A., Itatani, N., & Klump, G. M. (2010). A comparative view on the perception of mistuning: Constraints of the auditory periphery. In E. Lopez-Poveda, A. R. Palmer, & R. Meddis (Eds.), *The neurophysiological bases of auditory perception* (pp. 465–475). Springer New York.

Klink, K. B., & Klump, G. M. (2004). Duration discrimination in the mouse (*Mus musculus*). *Journal of Comparative Physiology. A*, 190, 1039–1046. https://doi.org/10.1007/s00359-004-0561-0

Kohlrausch, A. G., & Sander, A. (1995). Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets. *The Journal of the Acoustical Society of America*, 97, 1817–1829. https://doi.org/10.1121/1.413097

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, 3615–3627. https://doi.org/10.1121/1.2799482

Lauer, A. M., Mollis, M., & Leek, M. R. (2009). Discrimination of time-reversed harmonic complexes by normal-hearing and hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology*, 10, 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. In D. Pressnitzer, A. de Cheveigné, S. McAdams, & L. Collet (Eds.), *Auditory signal processing* (pp. 470–476). Springer New York.

Müller, M. (1996). The cochlear place-frequency map of the adult and developing Mongolian gerbil. *Hearing Research*, 94, 148–156. https://doi.org/10.1016/0378-5955(95)00230-8

National Institute of Mental Health. (2002). *Methods and Welfare Considerations in Behavioral Research with Animals: Report of a National Institutes of Health Workshop*. NIH Publication No. 02-5083. U.S. Government Printing Office.

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, 2173–2186. 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, 522–539. https://doi.org/10.1152/jn.00776.2006

Oxenham, A. J. (2018). How we hear: The perception and neural coding of sound. *Annual Review of Psychology*, 69, 27–50. https://doi.org/10.1146/annurev-psych-122216-011635

Oxenham, A. J., & Dau, T. (2001). Towards a measure of auditory-filter phase response. *The Journal of the Acoustical Society of America*, 110, 3169–3178. https://doi.org/10.1121/1.1414706

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, 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, 169–175. https://doi.org/10.1016/0006-8993(78)90320-7

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience*, 21, 227–277. https://doi.org/10.1146/annurev.neuro.21.1.227

Plassmann, W., Peetz, W., & Schmidt, M. (1987). The cochlea in gerbille rodents. *Brain, Behavior and Evolution*, 30, 82–102. https://doi.org/10.1159/000118639

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, 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, 2281–2298. https://doi.org/10.1121/1.1318898

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, 769–793. https://doi.org/10.1152/jn.1967.30.4.769

Ryan, A. F. (1976). Hearing sensitivity of the Mongolian gerbil, *Meriones unguiculatus*. *The Journal of the Acoustical Society of America*, 59, 1222–1226. https://doi.org/10.1121/1.380961

Schnupp, J. W., Garcia-Lazaro, J. A., & Lesica, N. A. (2015). Polysteposity in the gerbil inferior colliculus: Local clustering rather than a gradient map. *Frontiers in Neural Circuits*, 9, 37. https://doi.org/10.3389/fncir.2015.00037

Schreiner, C. E., & Langner, G. (1988). Periodicity coding in the inferior colliculus of the cat. II. Topographical Organization. *Journal of Neurophysiology*, 60, 1823–1840. https://doi.org/10.1152/jn.1988.60.6.1823

Schroeder, M. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (Corresp.). *IEEE Transactions on Information Theory*, 16, 85–89. https://doi.org/10.1109/TIT.1970.1054411

Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2010). Otoacoustic estimation of cochlear tuning: Validation in the chinchilla. *Journal of the Association for Research in Otolaryngology*, 11, 343–365. https://doi.org/10.1007/s10162-010-0217-4

Sinex, D. G., López, D. E., & Warr, W. B. (2001). Electrophysiological responses of cochlear root neurons. *Hearing Research*, 158, 28–38. https://doi.org/10.1016/S0378-5955(01)00293-3

Smith, B. K., Sieben, U. K., Kohlrausch, A., & Schroeder, M. R. (1986). Phase effects in masking related to dispersion in the inner ear. *The Journal of the Acoustical Society of America*, 80, 1631–1637. https://doi.org/10.1121/1.394327

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, 294–306. https://doi.org/10.1121/1.1580813

van Rossum, M. C. W. (2001). A novel spike distance. *Neural Computation*, 13, 751–763. https://doi.org/10.1162/089976601300014321

Versteegh, C., Meenderink, S. W. F., & van der Heijden, M. (2011). Response characteristics in the apex of the gerbil cochlea studied through auditory nerve recordings. *Journal of the Association for Research in Otolaryngology*, 12, 301–316.

Walton, J. P. (2010). Timing is everything: Temporal processing deficits in the aged auditory brainstem. *Hearing Research*, 264, 63–69. https://doi.org/10.1016/j.heares.2010.03.002

Wright, A., Davis, A., Bredberg, G., Ülehlová, L., & Spencer, H. (1987). Hair cell distributions in the normal human cochlea: A report of a European working group. *Acta Oto-Laryngologica*, 104, 15–24. https://doi.org/10.3109/00016488709124972

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.