

Frequency difference limens of pure tones and harmonics within complex stimuli in Mongolian gerbils and humans

Astrid Klinge^{a)} and Georg M. Klump

*Animal Physiology and Behavior Group, Department of Biology and Environmental Sciences,
Carl-von-Ossietzky University Oldenburg, Postfach 2503, 26111 Oldenburg, Germany*

(Received 7 March 2008; revised 10 October 2008; accepted 16 October 2008)

Frequency difference limens (FDLs) for pure tones between 200 and 6400 Hz and for the first, the second, or the eighth harmonic of an 800 Hz complex in four Mongolian gerbils (*Meriones unguiculatus*) were determined using a Go/NoGo-procedure. The 12 harmonics of the complex started either in sine phase or at a random phase. Gerbils showed very high pure tone FDLs ranging from 17.1% Weber fraction (200 Hz) to 6.7% (6400 Hz). They performed much better in detecting mistuning of a harmonic in the complex in the sine phase condition with FDLs decreasing from 0.07% for the first harmonic to 0.02% for the eighth harmonic. FDLs were about one order of magnitude higher when temporal cues were degraded by randomizing the starting phase of every component in the harmonic complex for every stimulus. These results are strikingly different from those obtained in four human subjects who needed about four times higher frequency shifts than gerbils for detecting a mistuned component in a sine phase complex and showed similar detection of mistuning in the random phase condition. The results are discussed in relation to possible processing mechanisms for pure tone frequency discrimination and for detecting mistuning in harmonic complex stimuli. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3021315]

PACS number(s): 43.66.Fe, 43.66.Gf, 43.66.Nm, 43.66.Hg [WPS]

Pages: 304–314

I. INTRODUCTION

The natural environment is acoustically complex. The auditory system receives acoustic information from many different sound sources that often overlap in time and in frequency. The auditory system has the capability to segregate these sound sources and to group together sounds that belong to one source. Many of the natural sounds are not simple pure tones but are harmonic complexes composed of tonal components with a frequency that is an integer multiple of a fundamental frequency f_0 . Such harmonic complexes are produced by the human voice, musical instruments, and animal calls, but can even be generated incidentally by mechanical action of water or wind. The harmonic complex is generally perceived as one sound with a pitch corresponding to the fundamental frequency. Thus, harmonicity is a strong cue to group together frequencies aiding the segregation of sounds from different sources and the grouping of sounds from a single source to a perceived auditory object.

If a component of a harmonic complex is shifted in frequency, the percept of the harmonic complex changes for small frequency shifts until the mistuned component falls out of the harmonic series producing a separate auditory object at a large frequency shift. The “mistuned” harmonic paradigm is one way of studying the influence of harmonicity on the segregation of concurrent auditory objects (e.g., Moore *et al.* 1984, 1985, 1986; Lin and Hartmann, 1998). In such an experiment, a harmonic complex serves as a reference stimulus and in the inharmonic test stimulus one of the components of the harmonic complex is shifted in frequency.

Moore *et al.* (1985) reported that the cues for detecting the inharmonicity differed for the lower and higher harmonics. Mistuned harmonics of low frequencies that are resolved by the auditory system “stand out” as a separate pure tone whereas harmonics with high unresolved frequencies are heard as a “beating” of the stimulus. These results indicate that in a harmonic complex with a mistuned component not only changes in the spectral domain but also changes in the temporal structure can be used to detect inharmonicity.

In this paper, we evaluate the perception of mistuned components in an otherwise harmonic complex in the Mongolian gerbil and compare these results to human data. Mongolian gerbils are a widely used animal model for studying auditory processing mechanisms in mammals since their hearing sensitivity in the low-frequency region is similar to that of humans. We measured the gerbil’s ability to detect a frequency shift in three behavioral experiments. In the first experiment, we obtained frequency difference limens (FDLs) for pure tones at different frequencies that were used as a reference for the two subsequent experiments. In the second and third experiment, FDLs for mistuned components of a harmonic complex stimulus were measured. The harmonic complex in the second experiment consisted of components that all started in sine phase. This resulted in a periodic temporal pattern that may provide a reference for detecting a change caused by the mistuning of a harmonic. If gerbils are especially sensitive to such temporal cues, we expect them to perform much better in detecting a frequency shift leading to the inharmonicity in a complex stimulus than in frequency discrimination of pure tones as has been found in birds (e.g., Lohr and Dooling, 1998). By randomizing the phase of every component in the harmonic complex in every presented stimulus during the third experiment, we provided a refer-

^{a)}Author to whom correspondence should be addressed. Electronic mail: astrid.klinge@uni-oldenburg.de

ence with a constant spectral composition but a varying temporal structure. We thereby made a change in the temporal pattern less useful as a cue. The comparison of the thresholds for detecting a frequency shift in the three conditions will provide insight into the mechanisms of processing of harmonic complexes that affects the formation of auditory objects.

II. MATERIALS AND METHODS

A. Animal subjects

Four adult Mongolian gerbils (*Meriones unguiculatus*)—three females, one male—were trained to serve as subjects in the experiments. For one of the females, we only obtained hearing thresholds and FDLs of mistuned harmonics in a complex with all harmonics starting in sine phase. During the time of testing, the animals were between 15 and 27 months old. All four gerbils had normal hearing within the frequency range of 100 Hz–8 kHz (see below for further explanation). Animals were housed in individual cages (Type III, $42 \times 26 \times 15$ cm³) and had unlimited access to water. The daily amount of food was restricted to maintain a weight of approximately 80% of their *ad libitum* weight.

The care and treatment of the gerbils were in accordance with the procedures of animal experimentation approved by the Government of Lower Saxony, Germany. All procedures were performed in compliance with the NIH Guide on Methods and Welfare Considerations in Behavioural Research with Animals (NIH Publication No. 02-5083).

B. Apparatus and stimulus generation

Experiments were conducted in a single-walled sound attenuating booth (Industrial Acoustics, Type IAC 401-A). The interior of the chamber was lined with a 15 cm thick layer of sound absorbing foam with a pyramidal surface (Illbruck Illtec Pyramide 100/50, Illbruck Illtec PLANO Type 50/0). The reverberation time T_{60} of a broadband white noise measured through a microphone placed at the location where the head of the gerbil would normally be during the experiment was 12 ms and thus can still be referred to as anechoic.

The experimental cage was mounted on a rack in the middle of the chamber. The rack was built out of 5 mm thin metal bars that produced no substantial reflections. The cage was doughnut shaped and consisted of wire mesh (mesh size of 1 cm) on all sides. The floor of the cage was covered with fine wire mesh (mesh size of 1 mm) to provide for an even surface. Inside the cage, there was an elevated platform where the gerbils had to sit and wait for the test stimulus being played. Two light barriers were installed, one monitoring the pedestal and one the access to it. By forcing the gerbil to interrupt the light barriers in a specific sequence, we ensured that test stimuli were only presented if the gerbil was facing the loudspeaker. The loudspeaker (Canton Plus XS, frequency range: 150 Hz–21 kHz) was mounted about 30 cm in front of the elevated platform at 0° azimuth and 0° elevation in relation to the head of the gerbil. For correct responses, food rewards (20 mg round pellets, Bioserve Dustless Precision Pellets Formula F0163-J50) were dispensed from a custom-built feeder. It was connected via a

flexible tube to a food tray mounted close to the platform. A red feeder light emitting diode (LED) served as an additional reinforcer. Custom software controlled feeder and feeder light, registered switches of the light barriers, and generated and delivered the stimuli. Experiments took place with no visible light in the chamber. The sessions could be monitored by a charge coupled device camera and infrared LEDs.

All stimuli were generated using a Linux workstation with an AMD processor and an RME sound card (Hammerfall DSP Multiface II). The analog signal from the sound card (sample frequency of 44.1 kHz) was delivered to a manual attenuator (Kenwood type RA-920A, presetting of overall signal level) and from there to a programmable attenuator (PA5 from Tucker Davis Technologies System 3, for computer controlled variations in signal level). The signal was then passed through an acoustic isolation transformer (1:1) to an amplifier (Rotel High Current 8 Channel Power Amplifier RMB-1048) and from there to the free-field loudspeaker. Stimuli were monitored on an oscilloscope (Conrad Voltcraft Type 630-2).

The setup was calibrated before every testing day using a spectrum analyzer (Stanford Research SR780) in conjunction with a G.R.A.S. Type 21SP Acoustic Transducer (microphone Type 40AF plus preamplifier) which was permanently mounted next to the pedestal.

We conducted three experiments with three different stimuli. In the first experiment, we obtained FDLs for pure tones. Test frequencies were 200, 400, 800, 1200, 1600, 3200, 6000, and 6400 Hz. Tone duration was 400 ms including a 25 ms Hanning ramp at stimulus onset and offset. The overall sound-pressure level (SPL) of 60 dB was roved by ± 3 dB to eliminate any possible level cue for the gerbils.

In the second and third experiment, we obtained FDLs for mistuned components in a harmonic complex. The harmonic complex consisted of the first 12 harmonics with a fundamental frequency f_0 of 800 Hz and a SPL of 60 dB per component (with levels adjusted to the frequency response of the loudspeaker). The SPL of the overall complex was varied randomly by ± 3 dB. As in the first experiment, the duration of the stimulus was 400 ms including a 25 ms Hanning ramp at onset and offset. In the reference stimuli, all components were harmonically related whereas in the test stimuli one of the harmonics was mistuned. For the second experiment, all 12 components of the complex started in sine phase [0°, Fig. 1(a)]. This resulted in reference stimuli that had exactly the same peaky waveform from one presentation to the next and the waveform was only different if a test stimulus with a mistuned harmonic was presented [Figs. 1(b) and 1(c)]. For the third experiment, all components started at a random phase which resulted in reference stimuli [Fig. 1(d)] and test stimuli [Fig. 1(e)] with a different temporal pattern for every stimulus presentation. We determined FDLs for frequency shifts of the first (800 Hz), the second (1600 Hz), and the eighth (6400 Hz) harmonic.

C. Procedure

Animals were trained using operant conditioning with positive reinforcement (food pellets) in a Go/NoGo para-

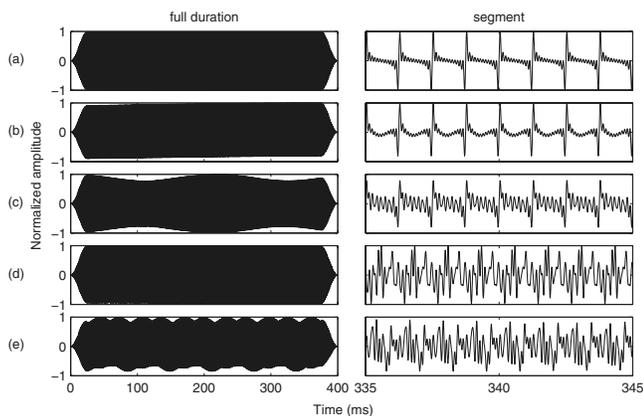


FIG. 1. Complex stimulus consisting of 12 harmonics and a fundamental frequency of 800 Hz. Left column shows complete stimulus length of 400 ms, right column a close up to compare the changes in the fine structure. [(a)–(c)] harmonic complex stimulus with all harmonics starting at sine phase with (a) no mistuning, (b) first harmonic mistuned by 0.07% (0.6 Hz threshold value), and (c) eighth harmonic mistuned by 0.07% (4.7 Hz); [(d)–(e)] harmonic complex stimulus with all harmonics starting at random phase with (d) no mistuning and (e) first harmonic mistuned by 2% (15.8 Hz threshold value). Note the differences in fine structure of every stimulus.

digm. During the session, we continuously repeated a reference stimulus every 1.3 s. The stimulus was either a pure tone (experiment 1) or a harmonic complex (experiments 2 and 3). To initialize a trial, the gerbil had to jump onto the pedestal and interrupt the light barrier. After a randomly chosen waiting interval between 1 and 7 s, the test stimulus was played instead of the reference stimulus. The test stimulus could either be the same as the reference stimulus during a catch trial or be different from the reference stimulus. For the first experiment, the test stimulus was a pure tone with a higher frequency than the reference stimulus. In the second and third experiment, the test stimulus was an inharmonic complex with one of the harmonics shifted upwards in frequency.

A correct response after a test stimulus (“Hit”) was rewarded by switching on the feeder LED for 3 s and delivering a food pellet with a probability of 80%. Sessions consisted of a block of ten warm-up trials, followed by ten blocks of ten trials each. Within each block, three catch trials and seven different test trials were presented in a randomized order. Frequency increments in test trials consisted of fixed steps of a preselected Weber fraction. The step size was adjusted for the FDL measurements at the different harmonics. To provide at least some salient frequency differences to the animals, the step size for the two largest frequency increments that were not included in determining thresholds was double that of the step size used for determining thresholds. For example, a possible step size of 0.00005 Weber fraction around the threshold value (1.1 Hz in the 6400 Hz component) in the sine phase condition would result in frequency shifts of 0.32, 0.64, 0.96, 1.28, 1.60, 2.24, and 2.88 Hz for the seven test stimuli. In the block of warm-up trials, only the largest frequency increment was presented. The performance during these trials was not included in the data analysis.

D. Testing of human subjects

We obtained FDLs from four human subjects listening to the same harmonic complex stimuli that were generated by similar hardware as in the experiments with the gerbils. Subjects included the first author and three students, who took part in a practical course. All four subjects had normal hearing. Informed written consent was obtained from each subject. Except for the first author, the human subjects had no prior experience in psychoacoustic tasks. The students got course credits for completing the experiments. The experiments were conducted in a single-walled sound attenuating chamber (IAC Mini 250). The stimuli were presented via headphones (Sennheiser HDA 200). The headphones were calibrated using an artificial ear (Brüel&Kjær Type 4153) mounted on a microphone (Brüel&Kjær Type 4134) and a measuring amplifier (Brüel&Kjær Type 2610). All harmonics were presented at 60 dB SPL. The procedure being used in the gerbil experiments was adapted for humans, replacing the light barriers with a two-button response box. The human subjects had to press one of the buttons to initiate a trial and the other button to respond to a difference between reference and test stimulus. Feedback was given at correct responses by switching on a LED on the response box for 2 s.

E. Data analysis

A session was accepted as being valid if the subject responded correctly to at least 80% of the two test stimuli with the largest frequency difference and their false alarm rate did not exceed 20%. A psychometric function was constructed relating d' to the frequency shift in the test trials. The threshold was determined by linearly interpolating between adjacent points of the psychometric function as the amount of frequency shift resulting in a d' of 1.8 (Green and Swets, 1966). Two consecutive valid sessions in which thresholds differed no more than 1% Weber fraction for experiment 1 and experiment 3 and 0.1% Weber fraction for experiment 2 were combined to calculate the final FDL.

To exclude any training effects, we randomized the presentation order of every threshold for every animal within one experiment and repeated the first threshold after finishing each experiment. If the repeated threshold did not differ more than 3% (experiment 1), 1% (experiment 3), or 0.1% (experiment 2), respectively, from the threshold obtained the first time, then we assumed that no training effect affected the results. If it differed by more than that, the second threshold had to be repeated, then the third and so on until the repetition matched the original results. Data were always taken from the last measurement. The order in which the experiments were conducted was randomized and different for each animal. For a statistical analysis we used the statistical software package STATISTICA (StatSoft®, Version 7.0). Throughout the text, average results are reported as mean \pm two times the standard error (mean \pm 2SE). All p -values are two tailed unless stated otherwise.

F. Prior testing of the hearing sensitivity of gerbils

Prior to the frequency discrimination experiments, we tested the hearing sensitivity of all four gerbils at the fre-

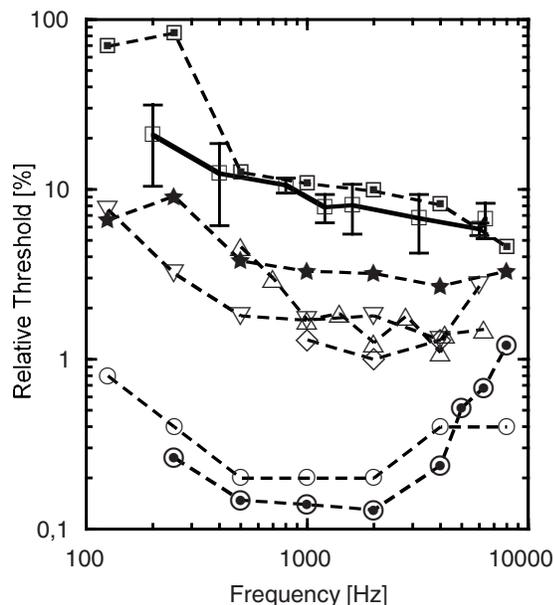


FIG. 2. Relative thresholds of pure tone frequency discrimination of the gerbils of the present study (open squares, continuous line, error bars = 2SEs), in comparison to literature data: gerbil (filled squares), African monkey (*Cercopithecus mitis*, *Cercocebus albigena*, stars), and human (open circles): Sinnott *et al.* (1992); human (filled circles): Moore (1973); budgerigar (open diamond): Dent *et al.* (2000); starling (open standing triangles): Klump *et al.* (2000); and pigeon (open hanging triangles): Sinnott *et al.* (1980).

frequencies of 100, 500, 1000, 2000, 4000, and 8000 Hz. Testing took place in the same chamber with the same procedure as was used for the frequency discrimination experiments. No stimulus was played except in test trials in which a single 800 ms pure tone was presented as the test signal. During a catch trial, no signal was presented. The method of constant stimuli was used with a step size in signal level of 3 dB. A session was accepted if at least 80% of responses to the two stimuli with the highest level were “Hits” and no more than 20% false alarms were reported. The hearing threshold was defined as the SPL at a d' of 1.8 (Green and Swets, 1966). Two consecutive valid sessions in which thresholds differed no more than 3 dB were combined to calculate the final hearing threshold based on 20 responses from 20 trials at each level and 60 catch trials. The presentation order of all tested frequencies was randomized for every animal and the first threshold was repeated after finishing the test series of the experiment to exclude any training effects. If the repeated

TABLE I. FDLs of pure tones. Individual data of three gerbils and their mean with 2SE in Hz.

Frequency (Hz)	Lv	Sa	Th	Mean \pm 2SE
200	45	22	58	42 \pm 21
400	67	25	56	49 \pm 25
800	84	77	72	78 \pm 7
1200	84	112	86	94 \pm 18
1600	87	147	153	129 \pm 42
3200	158	196	295	216 \pm 81
6000	377	327	345	350 \pm 29
6400	383	528	374	428 \pm 100

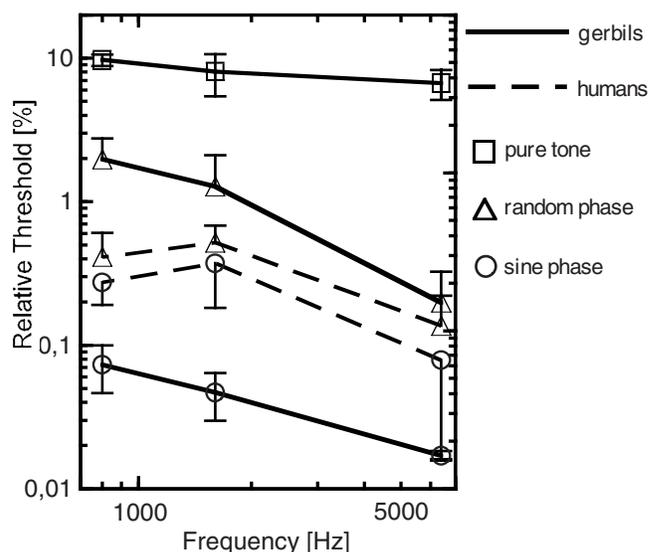


FIG. 3. Mean thresholds for gerbils and human subjects tested on the frequency discrimination ability of pure tones (only gerbils, squares) and on the frequency discrimination ability of the mistuned 800, 1600, and the 6400 Hz harmonic within a complex with all harmonics starting at random phase (triangles) and within a complex with all harmonics starting at sine phase (circles). Error bars represent 2 SE.

threshold differed more than 3 dB from the threshold obtained the first time, then the next threshold of the series was repeated. The range of tested frequencies with the best hearing was between 1 kHz (-1.0 ± 4.1 dB SPL, mean \pm standard deviation) and 8 kHz (-8.9 ± 1.5 dB SPL) with an average of -5.8 dB SPL. At the frequencies of 2 kHz and 4 kHz, we determined hearing sensitivity thresholds of -4.7 ± 2.7 dB SPL and -8.4 ± 1.3 dB SPL, respectively. Below 1 kHz, the hearing sensitivity declined and thresholds increased up to 16.2 ± 3.7 dB SPL at a frequency of 500 Hz and 38.0 ± 1.3 dB SPL at 100 Hz. In the low-frequency range up to 1 kHz, the gerbils’ hearing did not differ much from the hearing sensitivity thresholds obtained by Ryan (1976). At frequencies from 1 kHz to 8 kHz we observed between 7 and 14 dB SPL lower hearing thresholds than Ryan (1976). All four gerbils could be considered as having a normal hearing.

III. RESULTS

A. Experiment 1: Pure tone frequency discrimination

Figure 2 shows the mean pure tone FDLs of the gerbils in percent Weber fraction (solid line with open squares) in relation to the reference frequency. The corresponding FDLs in Hertz are listed in Table I. The relative frequency shift necessary for the detection decreased from a Weber fraction of $20.8\% \pm 10.4\%$ (mean \pm 2SE) at 200 Hz to $6.7\% \pm 1.6\%$ at 6400 Hz. This corresponds to an absolute frequency difference of 41.7 ± 20.9 Hz at a reference frequency of 200 Hz. The absolute frequency shift necessary for detection increased gradually with increasing frequency of the reference signal up to 428.2 ± 100.2 Hz at 6400 Hz. A one-way repeated measure ANOVA revealed a significant effect of the reference frequency on the relative frequency shift at detec-

TABLE II. FDLs for three harmonics of an 800 Hz complex with all components starting at sine phase and for three harmonics of an 800 Hz complex with all components starting at random phase. Individual data of three and four gerbils, respectively, and of four human subjects and their mean with 2SE in Hz.

Gerbils						
Sine phase condition						
Harmonic number	Frequency of harmonic (Hz)	Lv	Sa	Th	Ls	mean \pm 2SE
1	800	0.51	0.60	0.36	0.87	0.59 \pm 0.21
2	1600	0.44	0.97	0.60	0.99	0.75 \pm 0.27
8	6400	1.14	1.03	1.18	1.01	1.09 \pm 0.08
Random phase condition						
Harmonic number	Frequency of harmonic (Hz)	Lv	Sa	Th		mean \pm 2SE
1	800	19.9	9.2	18.4		15.8 \pm 6.7
2	1600	33.4	16.6	11.4		20.5 \pm 13.3
8	6400	5.8	19.8	12.5		12.7 \pm 8.1
HUMANS						
Sine phase condition						
Harmonic number	Frequency of harmonic (Hz)	AK	AV	NG	AT	mean \pm 2SE
1	800	1.9	3.1	2.2	1.5	2.2 \pm 0.7
2	1600	7.6	6.5	1.5	8.2	6.0 \pm 3.0
8	6400	2.9	10.5	5.5	1.3	5.1 \pm 4.0
Random phase condition						
Harmonic number	Frequency of harmonic (Hz)	AK	AV	NG	AT	mean \pm 2SE
1	800	1.9	3.5	2.5	5.4	3.3 \pm 1.5
2	1600	9.8	4.7	8.9	10.5	8.3 \pm 2.6
8	6400	2.4	15.1	7.1	5.5	8.8 \pm 5.4

tion threshold ($p < 0.05$; $\eta^2 = 0.56$) as well as on the absolute frequency shift at detection threshold ($p < 0.05$; $\eta^2 = 0.95$). The *post hoc* pairwise comparisons revealed a significant difference between the mean thresholds for the absolute frequency discrimination determined for 800 and 6400 Hz reference tones (Tukey test, $p < 0.01$). None of the other *post hoc* comparisons revealed significant differences.

B. Experiment 2: Frequency difference limens of mistuned components in a sine phase harmonic complex

In this experiment, we presented a harmonic complex to the gerbils in which every component started in sine phase [0° , Fig. 1(a)]. Applying a frequency shift on one of the harmonics produced a change in the fine structure of the waveform over the duration of the stimulus and a slight amplitude modulation occurred [Figs. 1(b) and 1(c)]. This stimulus condition resulted in the lowest FDLs of all three experiments. Figure 3 shows the amount of mistuning at threshold expressed as the Weber fraction relative to the frequency of the respective harmonic. The corresponding absolute threshold values in hertz are shown in Table II. The Weber fraction necessary for detecting the mistuning decreased from 0.073% \pm 0.027% (mean \pm 2SE) at the 800 Hz harmonic to 0.017% \pm 0.001% at the 6400 Hz harmonic.

Harmonic number significantly affected the Weber fraction necessary for detection of the mistuning (one-way repeated measure ANOVA, $p < 0.05$, $\eta^2 = 0.81$). Pairwise *post hoc* tests revealed a significant decrease of the threshold Weber fraction between the data for the 800 and 6400 Hz harmonic (Tukey test, $p < 0.01$) and a trend for the decrease comparing the 1600 and 6400 Hz harmonic (Tukey test, $p = 0.08$). The corresponding absolute frequency shift significantly increased from 0.59 Hz \pm 0.21 Hz at the 800 Hz harmonic to 1.09 Hz \pm 0.08 Hz at the 6400 Hz harmonic [one-way repeated measure analysis of variance (ANOVA), $p < 0.05$, $\eta^2 = 0.69$, a *post hoc* Tukey test revealed a significant difference between the data for the 800 Hz and the 6400 Hz harmonic, $p < 0.05$].

C. Experiment 3: Frequency difference limens of mistuned components in a random phase harmonic complex

During the third experiment, we randomized the starting phase of every component of the 800 Hz complex for every stimulus we presented throughout the experiment in order to degrade temporal cues. Figure 3 shows the relative frequency shifts necessary for the detection of a mistuning. The corresponding absolute frequency shifts at threshold are listed in Table II. A one-way repeated measure ANOVA revealed a

significant difference in the relative frequency shifts at the detection threshold from $2.0\% \pm 0.8\%$ (mean ± 2 SSE) at the 800 Hz harmonic to $0.22\% \pm 0.16\%$ at the 6400 Hz harmonic ($p < 0.05$; $\eta^2 = 0.79$). However, pairwise *post hoc* comparisons only revealed a significant decrease in the relative frequency shift necessary for the detection of a mistuning from the 800 Hz harmonic to the 6400 Hz harmonic (Tukey test, $p < 0.05$). All other pairwise comparisons were not significant. A one-way repeated measure ANOVA revealed no significant effect of the harmonic on the absolute frequency shift in Hertz.

D. Comparison of all three experiments

Comparing the results of all three experiments with each other (Fig. 3), we observed a highly significant difference between the mean relative frequency shift at the detection threshold among the three different experimental conditions. A two-way repeated measure ANOVA with experimental conditions (stimulus type) and reference frequency as factors including the data from the three gerbils that completed all experiments resulted in values of $p < 0.001$ and $p < 0.05$, respectively. The stimulus type had a strong effect on the discrimination threshold (effect size $\eta^2 = 0.99$). The thresholds for the relative frequency shifts were the lowest for harmonics in sine phase stimuli, about one order of magnitude higher for harmonics in random phase stimuli, and about another order of magnitude higher for pure tones (all pairwise planned comparisons of the experimental condition within a reference frequency revealed significant differences, all $p < 0.05$). The second main effect reference frequency also showed a high effect size η^2 of 0.78. There was no significant interaction between the factors experimental condition and reference frequency (two-way repeated measure ANOVA, $p = 0.26$).

The performance of the gerbils during the sessions was different for the second and third experiments. Corresponding to the large difference in FDL, it was more difficult to obtain data in the random phase condition, than in the sine phase condition. In the sine phase condition, the gerbils needed on an average five sessions to obtain a threshold which was significantly less than in the random phase condition with an average of 14 sessions to obtain a threshold (paired t-test, $p < 0.05$). When the gerbil switched from the sine phase condition to the random phase condition, their false alarm rate was significantly increased (from 98 to 160 false alarms, averaged over five sessions of two gerbils, paired t-test, $p < 0.05$), whereas a slight decrease was observed when switching from random phase to sine phase (from 104 to 101 false alarms, averaged over five sessions of one gerbil, paired t-test, $p = 0.92$).

E. Comparison with human data

Human FDLs for frequency shifts in the harmonic complex were obtained by using the same procedure and stimuli as for the gerbils to be able to compare FDLs directly (Table II, Fig. 3). A two-way repeated measure ANOVA with experimental condition (sine phase and random phase) and reference frequency (800, 1600, and 6400 Hz) as factors re-

vealed no significant difference between the Weber fraction of the shift at threshold for sine phase and random phase stimuli ($p = 0.129$). However, the reference frequency as a main effect had a significant effect on the Weber fraction at threshold ($p < 0.01$, $\eta^2 = 0.79$). *Post hoc* tests revealed significant differences for the comparison of the data obtained for reference frequencies of 6400 and 1600 Hz (Tukey test, $p < 0.01$) and for reference frequencies of 6400 and 800 Hz (Tukey test, $p < 0.05$). No significant interactions between reference frequency and experimental condition were found. Neither experimental condition nor reference frequency had a significant effect on the absolute frequency differences necessary for the detection of the mistuning (two-way repeated measure ANOVA).

For sine phase stimuli, the Weber fraction of the shift at threshold was smaller in gerbils than in humans, i.e., gerbils were more sensitive for detecting the mistuning. This difference was significant for the 800 Hz harmonic (t-test, $p < 0.005$) and the 1600 Hz harmonic (t-test, $p < 0.02$), and there was a trend in the same direction for the 6400 Hz harmonic (t-test, $p = 0.096$). For random phase stimuli, the Weber fraction of the shift at threshold was larger in gerbils than in humans. This difference was significant for the 800 Hz harmonic (t-test, $p < 0.005$) and there was a trend for the 1600 Hz harmonic (t-test, $p = 0.091$).

IV. DISCUSSION

A. Pure tone frequency difference limens

Compared to the discrimination ability of other animal species (Fig. 2), the gerbil exhibits a poor ability to discriminate the frequency of pure tones (Fig. 2, see also Sinnott *et al.*, 1992), especially in the low-frequency region. Human data are shown from Moore (1973) and Sinnott *et al.* (1992) for comparison. The highest frequency discrimination acuity in humans being as low as 0.2% shift in frequency is found between 500 Hz and 2 kHz. Bird species vary in their ability to discriminate between pure tones of different frequencies. Pigeons, for example, exhibit fairly high FDLs at frequencies higher than 1 kHz that are, however, within the range of nonhuman mammals (Sinnott *et al.*, 1980). Klump *et al.* (2000) reported pure tone FDLs in European starlings from 4.6% for a frequency of 500 Hz to 1.5% for 6300 Hz. Studying budgerigars, Dent *et al.* (2000) found FDLs of around 1.3% at frequencies of 1, 2, and 4 kHz. The SPLs used for the experiments in the different studies were all around 60–65 dB SPL which compares well to the sensation level used in the present study.

Pure tone frequency discrimination can be achieved based on two different mechanisms. The first mechanism applies to the low-frequency region where phase locking of auditory nerve (AN) fibers to the stimulus waveform occurs. For these frequencies, the distribution and change of interspike intervals in the AN fibers can be exploited to estimate the change in frequency (e.g., Wever, 1949). The second mechanism mainly applies to the high frequency region where phase locking is no longer possible. It evaluates the place of excitation in the cochlea for frequency discrimination. Ideas for a place mechanism in which frequencies are

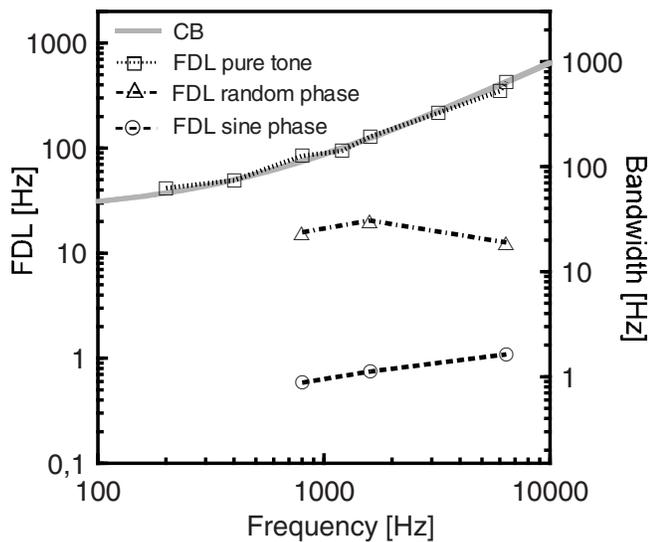


FIG. 4. Pure tone FDLs (black dotted line, squares), random phase FDLs (black dash-dotted line, triangles), and sine phase FDLs (black dashed line, circles) of gerbils are plotted in comparison to the cochlear place-frequency function (gray solid line). The CB function is derived from the general bandwidth function from Greenwood (1961) with parameters obtained from physiological data of gerbils (Müller, 1996).

represented tonotopically along the basilar membrane reach as far back as to the 17th century (Du Verney, 1683, c.f. Cheveigné, 2005). Later on, von Békésy (1960) and Zwicker (1970) proposed a place mechanism in which a frequency is represented as a certain excitation pattern along the basilar membrane. The detection of a change in frequency is based on changes in the excitation pattern.

Comparing a wide range of species, Fay (1992) favored a place mechanism in most animal species. He stated, based on analyses of cochlear place-frequency functions by Greenwood (1961, 1990), that perceptual units in the frequency domain such as the pure tone FDL, the critical bandwidth (CB), critical masking ratio bandwidth, and the psychophysical tuning curve can all be related to the spatial frequency map on the basilar membrane. In the Mongolian gerbil, a physiologically determined cochlear map (Müller, 1996) fitted by the cochlear place-frequency function (Greenwood, 1961) can be used to calculate the relation between the critical ratio bandwidth or the CB and the FDL (Kittel *et al.*, 2002). Figure 4 shows an estimate of the CB as a function of frequency (gray line) in comparison to the FDLs determined in this study. CBs were calculated from a cochlear place-frequency function that was fitted to the behavioral estimates of the CB (Kittel *et al.*, 2002). They range from 56 Hz at a center frequency of 200 Hz to 638 Hz at 6400 Hz. The function of the observed FDLs (black dotted line on top of the gray line) runs parallel to the CB function (i.e., it is offset by a scaling factor of 1.5) over the whole range of tested frequencies (Fig. 4, axes were adjusted accordingly). This parallel run of the CB function and the pure tone FDL function in the gerbil suggests that at least over the tested frequency range of 200–6400 Hz, a pure place mechanism might explain the observed FDLs.

What remains unexplained is the small scaling factor relating the FDL to the CB of the gerbil. Fay (1992) showed

that the best FDLs in humans are related to the CB by a factor of 50. He proposed that this scaling factor applies to most mammals and birds at least for frequencies above 1 kHz. However, Sinnott *et al.* (1992) argued against this scaling factor by presenting data from African monkeys (*Cercopithecus mitis*, *Cercocebus albigena*) that show much larger pure tone FDLs than humans in spite of having a CB that is similar to the human CB. The FDLs observed in monkeys can only be matched to the CB function by a scaling factor of 3 (see Sinnott *et al.*, 1992) which is more similar to the scaling factor of 1.5 found for the gerbil in the present study. Studying humans, Sek and Moore (1995) related different measures of frequency discrimination to the auditory filter bandwidth. They observed that the scaling factor is reduced considerably at frequencies for which humans have to rely on a place mechanism for frequency analysis. It deteriorated from about 18 at 4 kHz to about 5 at 8 kHz. This is similar to the range of scaling factors from about 7 to 17 that was found by Buus *et al.* (1995) fitting the FDL data from European starlings to the starling's CB function. These scaling factors, however, are much larger than the value found for the gerbil.

What may determine the value of the scaling factor and the value of the FDL that can vary considerably between species? Fay (1992) proposed that the largest absolute frequency discrimination acuity will be found in mammals with a poor high frequency hearing (thus perceiving a small frequency range) and a long cochlea. Frequency discrimination ability will be poor when the cochlea is short but the frequency range of sensitive hearing is wide. The gerbil cochlea is only 11 mm long (Müller, 1996) which is short in comparison to 34 mm reported for the human cochlea (Miller, 2007). Corresponding to the difference in length, the number of hair cells along the basilar membrane is reduced from about 3500 inner hair cells (IHCs) in the human cochlea (Wright *et al.*, 1987) to about 1400 IHCs in the gerbil cochlea (Plassmann *et al.*, 1987). Moreover, the hearing range of the gerbil extending from approximately 0.1–60 kHz (Ryan, 1976) is much larger than the human hearing range extending from about 0.01 to 16 kHz (ISO 389-7, 1996). Furthermore, the gerbil's larger frequency range is mapped onto more critical bands than the frequency range of humans (estimated on the basis of data from Kittel *et al.*, 2002 for gerbils, and Greenwood, 1961 and Moore and Glasberg, 1983 for humans). The distance on the cochlear map covered by a critical band in the gerbil is about 0.21 mm which is small compared to the distance of 1.15 mm estimated for humans (Fay, 1992). Thus, a specific shift in frequency relative to the reference will result in a smaller change of the spatial pattern of excitation in the cochlea involving a much smaller number of hair cells in gerbils compared to humans (a difference by a factor of 5–8). The difference between human and gerbil FDLs may thus be related to the higher number of hair cells in humans and encoding of frequencies by a larger number of AN fibers. This might result in a more fine-scaled frequency representation in the cochlea.

B. Frequency difference limens of mistuned components in a harmonic complex

The results of the “mistuned harmonic” experiments appeared in stark contrast to the pure tone FDLs. While the performance of the gerbils in pure tone frequency discrimination was unusually poor and about 20 times worse than the performance of humans, gerbils were about four times better than humans in detecting a mistuned harmonic in the sine phase condition. When gerbils detected a mistuned harmonic in a harmonic complex with all components starting at a random phase, FDLs were about one order of magnitude worse than FDLs obtained in the sine phase condition. This makes gerbils about two to four times less sensitive in reporting the mistuning compared to humans in the random phase condition.

1. Models for pitch perception and frequency discrimination

Two classes of models have been applied to explain pitch perception and discrimination of complex tone stimuli: (1) spectral pattern recognition models and (2) autocorrelation models. The spectral pattern recognition models (early models by Goldstein, 1973, Wightman, 1973, and Terhardt, 1974) assume that the frequency of every harmonic is determined either by place or by temporal mechanisms and then the best fitting fundamental frequency is estimated. An internal central processor or pattern recognizer decides if the frequency of a component of the complex is belonging to the harmonic series or not and has therefore been called a “harmonic sieve” by Duifhuis *et al.* (1982). One constraint of this type of model is, however, that the harmonics of the complex have to be resolved. Furthermore, the phase of components of complex stimuli is not taken into account as the pitch is extracted from the peripheral excitation pattern. A spectral pattern recognition model thus implies that the frequency selectivity of the basilar membrane has to be sufficiently good in order to reach low FDLs. To evaluate the involvement of place mechanisms in the frequency discrimination ability of mistuned harmonics in a harmonic complex, we plotted the FDLs of the random phase condition and of the sine phase condition of our study in relation to the place-frequency map in Fig. 4. If a place model would explain the FDLs for mistuned harmonics in complex stimuli at both condition, then the FDLs would be parallel to the cochlear place-frequency function determined by Greenwood (1961). The FDLs of mistuned harmonics at the random phase and the sine phase condition cannot be fitted to the cochlear map function by the same scaling factor at least over the tested frequency range which makes place mechanisms more unlikely to operate over the whole frequency range tested. Furthermore, we would expect much higher FDLs for mistuned harmonics in a harmonic complex in gerbils than in humans due to the gerbil’s poor pure tone FDLs, if we assume that the pattern recognition model applies. Our results do not support the prediction based on the spectral pattern recognition model.

Instead, the FDLs of mistuned harmonics in a complex with all harmonics starting at sine phase are two orders of

magnitude smaller than pure tone FDLs. This applies not only to the 6400 Hz harmonic that is unresolved in the gerbil auditory system but also for the 800 and the 1600 Hz harmonic that are resolved. Detecting an inharmonicity in the sine phase complex stimuli by the Mongolian gerbil therefore seems not to involve the analysis of the spectral pattern of a complex stimulus or the evaluation of temporal patterns within a single auditory filter but rather points to processing of temporal patterns beyond the stage of the basilar membrane.

Most of the temporal models for the processing of harmonic tone complexes rely on autocorrelation which is computed on the basis of the temporal envelope or temporal fine structure in all frequency channels. Such models were first introduced by Licklider (1951) and further developed by Meddis and Hewitt (1991a, 1991b), Meddis and O’Mard (1997), Patterson *et al.* (1995), and Denham (2005) using a summary autocorrelation function over all channels. The autocorrelation function suggested by Licklider (1951) is calculated on the basis of spike occurrences from a single neuron or a group of neurons and depends on the waveform envelope of unresolved harmonics to extract the pitch of a complex. A summary autocorrelation function combines the periodicity estimates derived from all frequency channels and therefore they do not depend only on unresolved components of a harmonic complex.

2. Temporal cues available for detecting a mistuned harmonic in a complex

Temporal cues that might be used for detecting mistuning of a component in a complex with all harmonics in sine phase could be a beating in the mistuned harmonic complex or a gradual phase shift over the duration of the stimulus resulting in a change of the fine structure of the ongoing stimulus. A beating is produced when the frequency-shifted harmonic in the mistuned harmonic complex is unresolved and interacts with the adjacent harmonics in the complex. Human subjects in the study by Moore *et al.* (1985) identified the mistuning of unresolved harmonics by hearing a kind of “beat.” This detection based on the amplitude modulation suggests that for humans temporal cues may play a greater role than spectral cues for detecting an unresolved mistuned harmonic in a complex stimulus. Temporal mechanisms for detecting a mistuned component in an otherwise harmonic complex in humans have also been suggested by Le Goff and Kohlrausch (2005). Their simulations showed that the excitation patterns for flat-spectrum and sloped-spectrum complex stimuli are less affected by mistuning the fundamental frequency component than the excitation pattern of the frequency-shifted pure tone. In a simulation of the envelope pattern of the output signal of the auditory filters, however, Le Goff and Kohlrausch (2005) showed that mistuning the fundamental frequency component in a sloped spectrum changed the temporal envelope of the signals in the auditory filters considerably. This points to the use of temporal mechanisms by humans even for resolved mistuned harmonics under certain circumstances.

Some authors proposed the evaluation of distortion products producing beating as a possible explanation for the

good performance in detecting a mistuned harmonic in a complex (Moore *et al.*, 1985; Pressnitzer and Patterson, 2001). Distortion products emerge at the point on the basilar membrane where the traveling waves of at least two nearby tones (with frequencies f_1 and f_2) overlap. Generally, the most prominent distortion product is the cubic $2f_1-f_2$ distortion which may under certain circumstances interact considerably with components of the presented stimulus on the basilar membrane. We estimated the possible interaction of the mistuned harmonic and the cubic distortion product which would develop at the place where the shifted harmonic has been. Taking into account measurements of distortion product otoacoustic emissions and neuronal responses by Faulstich and Kössl (1999) and Mom *et al.* (2001), we concluded that the SPL of the distortion product would be at least 20 dB lower than the level of the interacting mistuned harmonic. The estimated modulation depth due to the beating between the distortion product and the mistuned harmonic was around 1.1 dB. Since the gerbil's intensity discrimination threshold is at least 3 dB (Sinnott *et al.*, 1992), we conclude that the depth of modulation resulting from the beating of a distortion product with the mistuned component would not allow the gerbils to detect the beating and to use it as a detection cue.

A more likely temporal cue available to the gerbils for detecting the mistuning is an ongoing change in the fine structure over the duration of the mistuned harmonic complex. The change results from the gradual phase shift of the mistuned component compared to the other components. The peaky waveform of the sine phase stimulus may aid the gerbil in using this cue. In the random phase condition, however, only some of the randomly produced waveforms have a peaky temporal structure. Sensitivity to the phase or to changes of the phase in harmonic complex stimuli has been proposed in several physiological and psychophysical studies (e.g., Bilsen, 1973; Horst *et al.*, 1986, 1990; Moore and Glasberg, 1989; Simmons *et al.*, 1993). Phase relations might be detected on the basilar membrane when unresolved harmonics fall within one auditory filter. Phase relations for resolved harmonics might be preserved and detected in more central stages of the auditory system where they can be analyzed across channels. Hartmann (1988) showed that reversing the phase of the resolved fourth harmonic in comparison to the other components of a 200 Hz complex influenced the detection of a mistuning in relation to the duration of the stimulus. Since the mistuned component with the phase shift of 180° and the adjacent harmonics fall in different auditory filters, a processing of the phase relationship between the mistuned harmonic and the adjacent components is less likely to be analyzed on the basilar membrane. Thus, the phase effects observed by Hartmann (1988) suggest that the processing of these temporal cues in humans is not limited to cochlear within-channel mechanisms.

A further improvement for detecting a mistuned harmonic in the sine phase compared to the random phase stimulus could be explained by the stimulus statistics. In the sine phase condition, a constant temporal pattern is provided with every background stimulus that only changes during the presentation of a test stimulus. In the random phase stimulus

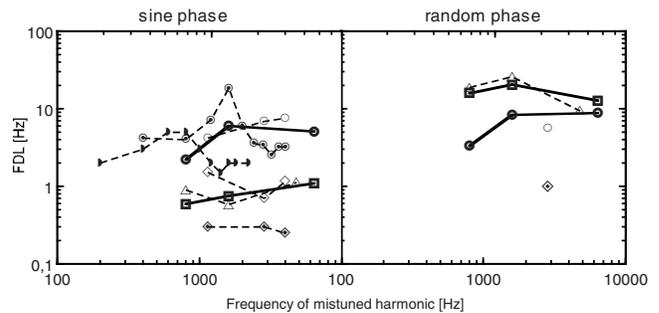


FIG. 5. Absolute thresholds in Hz of mistuned components of sine phase (left panel) and random phase (right panel) complex stimuli of gerbils (thick square, continuous line) and humans (thick circle, continuous line) determined in the present study in comparison to literature data. Zebra finch (filled diamond), budgerigar (open diamond) and human (open circle): Lohr and Dooling (1998); starling (open triangle): Klump *et al.* (2008); human (filled circle): Moore *et al.* (1985); human (half circle): Lee and Green (1994).

condition, we continuously varied the temporal fine structure and phase relations by randomizing the starting phase of every harmonic in every presented stimulus. Changing the temporal fine structure with every presented stimulus removed the possibility to compare temporal patterns of the harmonic complex sequentially. This high level of uncertainty in the stimuli might have resulted in the observed decline of FDLs in the random phase condition compared to the sine phase condition. The observation that gerbils needed more sessions to obtain a threshold in the random phase condition than in the sine phase condition and that they showed an increased false alarm rate when switching from the sine phase to the random phase condition also suggests that available cues in the random phase condition may have been less salient.

The amplitude modulation of the stimulus envelope resulting from mistuning could provide another possible cue for detecting a mistuned harmonic in a complex in the random phase condition. The depth of the amplitude modulation, however, has to be sufficiently large to be perceived by the gerbil. In the sine phase condition, the FDLs are too small to produce an exploitable modulation. Also in the random phase condition, the modulation spectrum of the mistuned complex stimulus at threshold value indicated that the modulation depth is not large enough to be detectable by the gerbil.

3. Comparison with other species

Figure 5 gives an overview of FDLs of mistuned components in harmonic complexes in several animal species and humans. Zebra finches and budgerigars exhibit very low FDLs for detecting mistuning of components in harmonic complexes in the sine phase condition (Lohr and Dooling, 1998) that were similar to the FDLs of gerbils determined in the present study. An increased FDL in the random phase condition compared to the sine phase condition could be shown for zebra finches in the study by Lohr and Dooling (1998). The European starling also exhibits discrimination thresholds that are similar to those found in gerbils (Klump *et al.*, 2008). The FDLs in the bird species were significantly lower compared to those of humans when all components of

the complex stimulus started in sine phase. The studies show that gerbils are not the only species that have such low FDLs for detecting a frequency shift in harmonic complexes in the sine phase condition. The difference in the salience of mistuning in the sine phase and the random phase conditions in gerbils and birds suggests that temporal fine structure of the stimuli may be of particular importance. Sinnott and Motteler (2001) found gerbils to be equally sensitive in detecting frequency changes in speech formants compared to detecting a frequency shift of a harmonic in the random phase experiment as determined in the present study. Threshold values for detecting frequency changes in speech formants were similarly decreased compared to pure tone FDLs in gerbils as is described here for detecting a mistuning of a harmonic in a random phase complex. Furthermore, the ratio between the gerbil and human formant FDLs in the speech discrimination task and between the gerbil and human FDL in the random phase condition was found to be similar. These comparisons could be interpreted as a hint that humans and gerbils may employ similar mechanisms in processing frequency changes in both experiments. However, it may well be that both species use different mechanisms resulting in a similar sensitivity for detecting the frequency shift.

Human subjects were not as sensitive as gerbils or zebra finches (Lohr and Dooling, 1998) to the loss of distinctive temporal features of the stimulus during the presentation of random phase harmonic complexes. In humans, the FDLs for mistuned harmonics in harmonic complexes with all components starting at random phase were never significantly different to the FDLs for mistuned harmonics in the sine phase condition (see present study or Lohr and Dooling, 1998). These results are supported by studies from Dooling *et al.* (2002) and Patterson (1987) who showed that humans cannot discriminate cosine phase complex stimuli from random phase stimuli at a fundamental frequency of 800 Hz. This points to a reduced importance of temporal fine structure as a cue in humans compared to gerbils and birds.

Which constraints may bias gerbils and birds to rely more and humans to rely less on temporal cues when detecting mistuning of components in a harmonic complex? The length of the cochlea may hint at a possible explanation. The Mongolian gerbil has a short cochlea compared to the human cochlea with much fewer hair cells covering a CB. Thus, the same change in frequency in a gerbil and a human cochlea results in a much smaller change of the spatial pattern of excitation in the gerbil than in humans. In order to compensate for this, the gerbil and possibly the birds (also having a short cochlea) might have switched to exploiting temporal cues in complex stimuli to detect small amounts of mistuning. However, by proposing the use of temporal fine structure as a cue the question remains why gerbils seem not to exploit any temporal cues for detecting frequency differences in pure tones (at least in the frequency range below 2 kHz for which this is suggested in humans, e.g., Moore, 1973). Since humans have a much longer cochlea than birds or the gerbil, the evolutionary pressure to use changes in the temporal fine structure as a cue in the processing of frequency shifts in harmonic complexes may have been much lower than in the other species discussed here.

ACKNOWLEDGMENTS

We thank Susanne Groß and Anne Vellage for helping with the experiments, Antje Tjarks, Nantje Gelszat, and Anne Vellage for being motivated subjects in the human experiments, Holger Dierker for technical support, Stephan Ewert for assisting in the reverberation measurement, and Stefan Strahl for helping out in computational problems. This study was supported by the DFG within the Sonderforschungsbereich SFB TR31 “The Active Auditory System” and the International Graduate School GRK 591 “Neurosensory Science and Systems.”

- Bilsen, F. A. (1973). “On the influence of the number and phase of harmonics on the perceptibility of the pitch of complex signals,” *Acustica* **28**, 60–65.
- Buus, S., Klump, G. M., Gleich, O., and Langemann, U. (1995). “An excitation-pattern model for the starling (*Sturnus vulgaris*),” *J. Acoust. Soc. Am.* **98**, 112–124.
- Cheveigné, A. (2005). “Pitch perception models,” in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. R. Fay, and A. N. Popper, (Springer, New York), pp. 169–233.
- Denham, S. L. (2005). “Pitch detection of dynamic iterated rippled noise by humans and a modified auditory model,” *BioSystems* **79**, 199–206.
- Dent, M. L., Dooling, R. J., and Pierce, A. L. (2000). “Frequency discrimination in budgerigars (*Melopsittacus undulatus*): Effects of tone duration and tonal context,” *J. Acoust. Soc. Am.* **107**, 2657–2664.
- Dooling, R. J., Leek, M. R., Gleich, O., and Dent, M. L. (2002). “Auditory temporal resolution in birds: Discrimination of harmonic complexes,” *J. Acoust. Soc. Am.* **112**, 748–759.
- Duifhuis, H., Willems, L. F., and Sluyter, R. J. (1982). “Measurement of pitch in speech: An implementation of Goldstein’s theory of pitch perception,” *J. Acoust. Soc. Am.* **71**, 1568–1580.
- Faulstich, M., and Kössl, M. (1999). “Neuronal response to cochlear distortion products in the anteroventral cochlear nucleus of the gerbil,” *J. Acoust. Soc. Am.* **105**, 491–502.
- Fay, R. R. (1992). “Structure and function in sound discrimination among vertebrates,” in *The Evolutionary Biology of Hearing*, edited by D. Webster, R. R. Fay, and A. Popper (Springer, New York), pp. 229–263.
- Goldstein, J. L. (1973). “An optimum processor theory for the central formation of the pitch of complex tones,” *J. Acoust. Soc. Am.* **54**, 1496–1516.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Greenwood, D. D. (1961). “Critical bandwidth and the frequency coordinates of the basilar membrane,” *J. Acoust. Soc. Am.* **33**, 1344–1356.
- Greenwood, D. D. (1990). “A cochlear frequency-position function for several species—29 years later,” *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Hartmann, W. M. (1988). “Pitch perception and the segregation and integration of auditory entities,” in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 623–645.
- Horst, J. W., Javel, E., and Farley, G. R. (1986). “Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms,” *J. Acoust. Soc. Am.* **79**, 398–416.
- Horst, J. W., Javel, E., and Farley, G. R. (1990). “Coding of spectral fine structure in the auditory nerve. II. Level-dependent nonlinear responses,” *J. Acoust. Soc. Am.* **88**, 2656–2681.
- ISO 389-7, (1996). “Acoustics—Reference zero for the calibration of audiometric equipment—Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions.”
- Kittel, M., Wagner, E., and Klump, G. M. (2002). “An estimate of the auditory-filter bandwidth in the Mongolian gerbil,” *Hear. Res.* **164**, 69–76.
- Klump, G. M., Groß, S., and Itatani, N. (2008). “Processing of a mistuned harmonic by the European starling auditory system,” 31st MidWinter Research Meeting [J. Assoc. Res. Otolaryngol. **31**, 139].
- Klump, G. M., Langemann, U., and Gleich, O. (2000). “The European starling as a model for understanding perceptual mechanisms,” in *Auditory Worlds: Sensory Analysis and Perception in Animals and Man*, edited by G. A. Manley, H. Fastl, M. Kössl, H. Oeckinghaus, and G. M. Klump (Wiley-VCH, Weinheim), pp. 193–211.
- Le Goff, N., and Kohlrausch, A. (2005). “The use of temporal cues for frequency discrimination of the fundamental component in a complex

- tone," Proceedings of the Forum Acusticum, Budapest, pp. 1503–1507.
- Lee, J., and Green, D. M. (1994). "Detection of a mistuned component in a harmonic complex," *J. Acoust. Soc. Am.* **96**, 716–725.
- Licklider, J. C. (1951). "A duplex theory of pitch perception," *Experientia* **7**, 128–134.
- Lin, J.-Y., and Hartmann, W. M. (1998). "The pitch of a mistuned harmonic: Evidence for a template model," *J. Acoust. Soc. Am.* **103**, 2608–2617.
- Lohr, B., and Dooling, R. J. (1998). "Detection of changes in timbre and harmonicity in complex sounds by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*)," *J. Comp. Psychol.* **112**, 36–47.
- Meddis, R., and Hewitt, M. J. (1991a). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," *J. Acoust. Soc. Am.* **89**, 2866–2882.
- Meddis, R., and Hewitt, M. J. (1991b). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity," *J. Acoust. Soc. Am.* **89**, 2883–2894.
- Meddis, R., and O'Mard, L. (1997). "A unitary model of pitch perception," *J. Acoust. Soc. Am.* **102**, 1811–1820.
- Miller, J. D. (2007). "Sex differences in the length of the organ of Corti in humans," *J. Acoust. Soc. Am.* **121**, EL151–EL155.
- Mom, T., Bonfils, P., Gilain, L., and Avan, P. (2001). "Origin of cubic difference tones generated by high-intensity stimuli: Effect of ischemia and auditory fatigue on the gerbil cochlea," *J. Acoust. Soc. Am.* **110**, 1477–1488.
- Moore, B. C. (1973). "Frequency difference limens for short-duration tones," *J. Acoust. Soc. Am.* **54**, 610–619.
- Moore, B. C., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Moore, B. C., and Glasberg, B. R. (1989). "Difference limens for phase in normal and hearing-impaired subjects," *J. Acoust. Soc. Am.* **86**, 1351–1365.
- Moore, B. C., Glasberg, B. R., and Peters, R. W. (1984). "Frequency and intensity difference limens for harmonics within complex tones," *J. Acoust. Soc. Am.* **75**, 550–561.
- Moore, B. C., Peters, R. W., and B. R. Glasberg (1985). "Thresholds for the detection of inharmonicity in complex tones," *J. Acoust. Soc. Am.* **77**, 1861–1867.
- Moore, B. C., Glasberg, B. R., and Peters, R. W. (1986). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," *J. Acoust. Soc. Am.* **80**, 479–483.
- Müller, M. (1996). "The cochlear place-frequency map of the adult and developing Mongolian gerbil," *Hear. Res.* **94**, 148–156.
- Patterson, R. D. (1987). "A pulse ribbon model of monaural phase perception," *J. Acoust. Soc. Am.* **82**, 1560–1586.
- Patterson, R. D., Allerhand, M. H., and Giguere, C. (1995). "Time domain modelling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**, 1890–1894.
- Plassmann, W., Peetz, W., and Schmidt, M. (1987). "The cochlea in Gerbilline rodents," *Brain Behav. Evol.* **30**, 82–101.
- Pressnitzer, D. and Patterson, R. D. (2001). "Distortion products and the perceived pitch of harmonic complex tones," in *Physiological and Psychophysical Bases of Auditory Function*, edited by D. J. Breebart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonoven (Shaker, Maastricht), pp. 97–104.
- Ryan, A. (1976). "Hearing sensitivity of the Mongolian gerbil, *Meriones unguiculatis*," *J. Acoust. Soc. Am.* **59**, 1222–1226.
- Sek, A., and Moore, B. C. J. (1995). "Frequency discrimination as a function of frequency, measured in several ways," *J. Acoust. Soc. Am.* **97**, 2479–2486.
- Simmons, A. M., Reese, G., and Ferragamo, M. (1993). "Periodicity extraction in the anuran auditory nerve II. Phase and temporal fine structure," *J. Acoust. Soc. Am.* **93**, 3374–3389.
- Sinnott, J. M., and Mosteller, K. W. (2001). "Comparative assessment of speech sound discrimination in the Mongolian gerbil," *J. Acoust. Soc. Am.* **110**, 1729–1732.
- Sinnott, J. M., Brown, C. H., and Brown, F. E. (1992). "Frequency and intensity discrimination in Mongolian gerbils, African monkeys and humans," *Hear. Res.* **59**, 205–212.
- Sinnott, J. M., Sachs, M. B., and Hienz, R. D. (1980). "Aspects of frequency discrimination in passerine birds and pigeons," *J. Comp. Physiol. Psychol.* **94**, 401–415.
- Terhardt, E. (1974). "Pitch, consonance and harmony," *J. Acoust. Soc. Am.* **55**, 1061–1069.
- Von Békésy, G. (1960). *Experiments in Hearing* (McGraw-Hill, New York).
- Wever, E. G. (1949). *Theory of Hearing* (Wiley, New York).
- Wightman, F. L. (1973). "The pattern-transformation model of pitch," *J. Acoust. Soc. Am.* **54**, 407–416.
- Wright, A., Davis, A., Bredberg, G., Ulehlova, L., and Spencer, H. (1987). "Hair cell distributions in the normal human cochlea," *Acta Oto-Laryngol., Suppl.* **444**, 1–48.
- Zwicker, E. (1970). "Masking and psychological excitation as consequences of the ear's frequency analysis," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).