Psychophysical and Physiological Evidence for Fast Binaural Processing

Ida Siveke,[1] **Stephan D. Ewert,**[2] **Benedikt Grothe,**[1,3] **and Lutz Wiegrebe**[1]

[1]Biocentre, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany, [2]Medizinische Physik, Fakultät V, Institut für Physik, Carl von Ossietzky Universität, 26111 Oldenburg, Germany, and [3]Bernstein Center for Computational Neuroscience Munich, 81377 Munich, Germany

The mammalian auditory system is the temporally most precise sensory modality: To localize low-frequency sounds in space, the binaural system can resolve time differences between the ears with microsecond precision. In contrast, the binaural system appears sluggish in tracking changing interaural time differences as they arise from a low-frequency sound source moving along the horizontal plane. For a combined psychophysical and electrophysiological approach, we created a binaural stimulus, called "Phasewarp," that can transmit rapid changes in interaural timing. Using this stimulus, the binaural performance in humans is significantly better than reported previously and comparable with the monaural performance revealed with amplitude-modulated stimuli. Parallel, electrophysiological recordings of binaural brainstem neurons in the gerbil show fast temporal processing of monaural and different types of binaural modulations. In a refined electrophysiological approach that was matched to the psychophysics, the seemingly faster binaural processing of the Phasewarp was confirmed. The current data provide both psychophysical and physiological evidence against a general, hard-wired binaural sluggishness and reconcile previous contradictions of electrophysiological and psychophysical estimates of temporal binaural performance.

Key words: binaural hearing; ITD; sound localization; amplitude modulation; binaural sluggishness; motion

Introduction

To localize low-frequency sounds in space, the mammalian auditory system relies on exquisitely precise estimation of time delays between the two ears. For low-frequency pure tones and noise, human psychophysical experiments show that interaural time differences (ITDs) as low as 10–20 μs can be resolved. In contrast to this extraordinary neural precision, the binaural system has been described as rather slow in following changes in ITDs as, for example, elicited by a low-frequency sound source moving in space. Previous experiments characterized binaural sluggishness by estimating the capability of the binaural system to detect masked spatially divergent signals (Grantham and Wightman, 1979; Kollmeier and Gilkey, 1990). Grantham and colleagues (Grantham and Wightman, 1978; Grantham, 1982) used a low-pass noise with time-varying ITD or time-varying interaural correlation to estimate the temporal precision of the binaural system. They found a binaural time constant of ~50–200 ms. The temporal resolution of monaural processing, based on amplitude-modulation detection, was quantified with time constants between 1.1 and 2.5 ms (Viemeister, 1979; Dau et al., 1999; Ewert and Dau, 2000; Kohlrausch et al., 2000). Together, these studies provided evidence that the binaural system is sluggish compared with the monaural system.

Electrophysiological studies showed that auditory midbrain responses are sensitive to modulations of interaural timing (Spitzer and Semple, 1991, 1993; Palmer et al., 1998; McAlpine et al., 2000) and modulations of interaural correlation (Joris et al., 2006). The latter study was the first that carefully studied the encoding of very high binaural modulation frequencies >100 Hz. Joris et al. (2006) showed that binaural neurons could lock to modulations of interaural correlation that are an order of magnitude faster than estimated from human psychophysical experiments (Grantham, 1982). Note that these studies used the same stimuli, the oscillating-correlation (Oscor) stimulus. Perceptually, this stimulus oscillates through stages of a single, focused spatial image, a completely diffuse image, and a blurred image created by anticorrelation. However, the Oscor does not feature modulations in ITD and thus does not convey the percept of auditory motion. Joris et al. (2006) argued that the fast binaural modulations presumably created in the superior olivary complex and transmitted to the inferior colliculus might not be processed fast enough by the auditory thalamus and cortex, creating the apparent binaural sluggishness found in psychophysical detection tasks.

A combination of human psychophysical experiments and electrophysiological recordings in the dorsal nucleus of the lateral lemniscus (DNLL) of a well established animal model of human sound localization, the Mongolian gerbil, is presented in the current study. The population of ITD-sensitive DNLL neurons was used because many DNLL neurons reflect ITD sensitivity of their inputs from the superior olivary complex (Seidl and Grothe,

2005; Kuwada et al., 2006; Siveke et al., 2006). A new stimulus, which creates strong and unambiguous auditory motion, the "Phasewarp," was created and the temporal resolution of monaural and binaural processing was directly compared. Unlike the Oscor, the Phasewarp produces a binaural modulation along the ITD axis, comparable with the ITD modulations produced by a noise sound source rotating around the head. The psychophysical data show that, with the Phasewarp, the auditory system can detect much faster binaural modulations than estimated previously. This psychophysical improvement is reflected in the responses of binaural neurons in the DNLL.

Parts of this work have been presented at the 14th International Symposium on Hearing, 2006.

Materials and Methods

Stimuli. To create a monaural modulation, independent Gaussian noises were multiplied with a sinusoidal modulator varying in amplitude between 0 and 2. The phase of the modulator was randomized over trials, but it was identical for the two ears. For low frequencies (less than ~10 Hz), this sinusoidal amplitude modulation (SAM) is perceived as a periodic change in loudness over time. For higher modulation frequencies around 60 Hz, the percept of roughness arises and for even higher modulation frequencies higher than ~125 Hz a faintly buzzing or whirring sound is perceived. The Oscor was generated according to Grantham (1982) starting with two independent noise samples. One was fed directly into the left ear. A copy of this noise sample was multiplied with a sine modulator. The other noise sample was multiplied with a cosine modulator. The two modulated waveforms were then added and fed to the right ear. The generation of the Oscor is illustrated in Figure 1 *A*. Perceptually, the Oscor stimulus oscillates through stages of a focused spatial image, a completely diffuse spatial image, and a blurred, semifocused image.

Figure 1. Illustration of the binaurally modulated stimuli: Oscor (*A*), Oscor01 (*B*), and Phasewarp (*C*). The left panels display the generation of the three stimuli. See text for a detailed description of the stimulus generation. The right panels show interaural cross-correlograms as an illustration of a binaural display of these stimuli. The degree of interaural correlation is grayscale coded as a function of time and ITD. Black represents the highest correlation. Whereas the Oscor and Oscor01 stimuli produce modulation only along the time axis, the Phasewarp stimulus produces binaural correlation modulation both along the time and ITD axis. mag, Magnitude; IFFT, inverse fast Fourier transformation.

As for the monaural amplitude modulation, this temporal course of the stimulus can be "followed" for modulation frequencies <10 Hz; for higher modulation frequencies, again the percept of roughness and faint buzzing arises. Another stimulus, the Oscor01, was generated also with two independent noise samples, one of them being fed directly to the left ear. A copy of this noise sample was multiplied with the square root of a raised-sine modulator, varying in amplitude between 0 and 1. The other noise sample was multiplied with the square root of the 180° phase shifted raised-sine modulator and the resulting waveforms were added and fed into the right ear. The generation of the Oscor01 is illustrated in Figure 1 *B*. In contrast to the Oscor stimulus, the Oscor01 stimulus does not oscillate through a blurred, semifocused image, which is produced by interaurally anticorrelated noise. Perceptually, the Oscor01 stimulus oscillates through a focused and completely diffuse spatial image. Changes in the quality of the percept for modulation frequencies >125 Hz are comparable with the Oscor. Phasewarp stimuli were generated in the frequency domain using a frequency independent magnitude and a random phase for the components of the spectrum of the stimulus for the left ear. For the spectrum of the right ear's stimulus, the phase components of the left ear were shifted along the frequency axis by an amount equal to the modulation fre-

quency. The generation of the Phasewarp stimulus is illustrated in Figure 1*C*. The interaural correlation of this stimulus oscillates not only as a function of time but also as a function of ITD. Perceptually, the Phasewarp produces the sensation of a rotation of a noisy sound source around the head. The velocity of this movement is reflected in the frequency of the modulation in ITD. Higher modulations produce faster movements until again the impression of roughness and faint buzzing arises for high modulation frequencies.
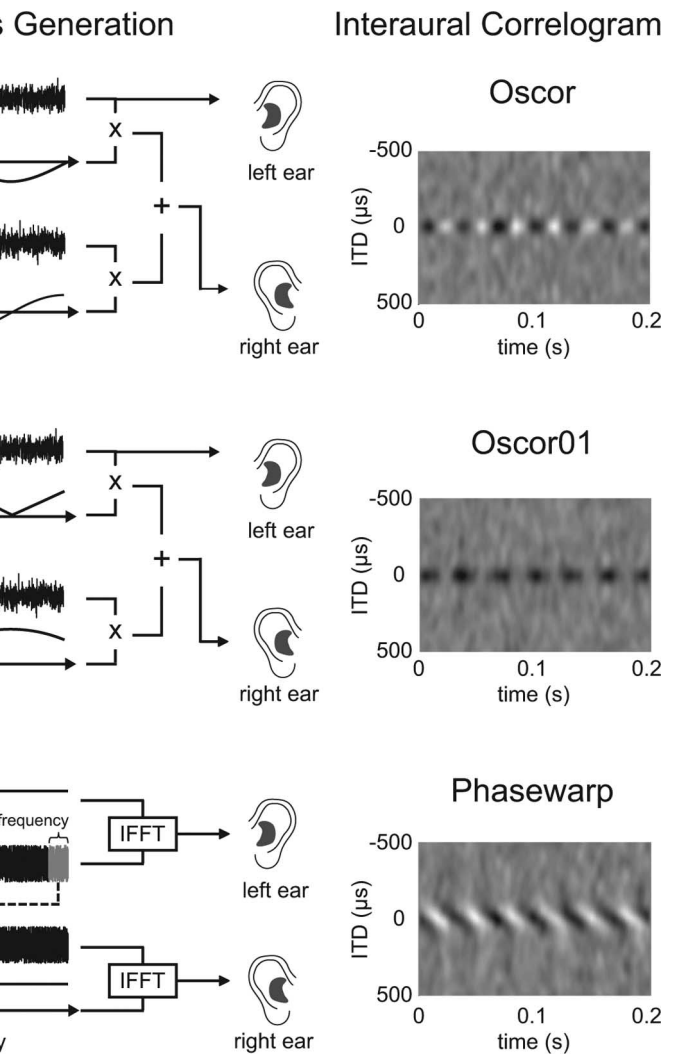
All stimuli were generated with modulation frequencies ranging from 2 to 512 Hz or from 8 to 1024 Hz in octave intervals. The rightmost column of Figure 1 shows the interaural correlation as a function of ITD and time for the three binaural stimuli, low-pass filtered at 5 kHz. Although the three stimuli share binaural modulations with the same modulation frequency of 8 Hz, the Phasewarp stimulus shows a pattern of correlation, which is modulated both along the ITD and the time axis. The pattern of the Oscor and Oscor01 stimulus is only modulated along the time axis. None of the binaural stimuli illustrated in Figure 1 produced any monaural amplitude modulation.

The standard (unmodulated) stimuli were samples of interaurally uncorrelated noise. In case of the Oscor01 measurements, a noise was used as standard in which the interaural correlation roved around the average long-
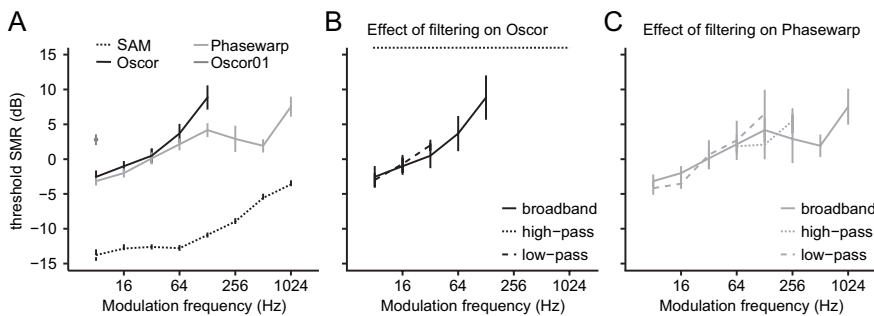
Figure 2. Average psychophysical sensitivity to monaural and three types of binaural modulations. Results with broadband stimulation are shown in **A**. The sensitivity, expressed in terms of the SMR at the modulation-detection threshold, is highest for the monaural modulations and worst for the Oscor01. The decrease of the sensitivity with increasing modulation frequency is very similar for the monaural modulations and the binaural modulations elicited by the Phasewarp stimuli. In contrast, with binaural modulations elicited by the Oscor stimuli, sensitivity decreases fast with increasing modulation frequency. With the Oscor01, binaural modulations could only be detected at the lowest tested modulation frequency (8 Hz). The effect of filtering on binaural modulation sensitivity is illustrated in **B** for the Oscor and in **C** for the Phasewarp stimuli. See text for details. Error bars indicate 1 SD.

term interaural correlation of the Oscor01, 0.5. The magnitude of the rove was ±0.25. This was done to prevent listeners from attending to the width of the binaural image, promoted by the interaural correlation, instead of the temporal modulation. In case of the Oscor, the average interaural correlation is 0 as for the interaurally uncorrelated noise standard.

In addition to the Phasewarp measurements with interaurally uncorrelated noise standard, a second standard, referred to as Phasewarp-equivalent correlation (PWEC) noise, was used in a psychophysical control experiment. The control experiment was required because the Phasewarp produces a relatively high interaural envelope correlation at the output of auditory filters when the modulation frequency is small with respect to the bandwidth of the auditory filter. This effect is particularly prominent at high characteristic frequencies, at which the equivalent rectangular bandwidth (ERB) of the filters is high, and at which phase-locking of the neuronal activity is progressively observed to the envelope and less to the fine structure of the stimulus. In combination, the Phasewarp results in an increased interaural correlation of the internal signals in the high-frequency auditory channels relative to the interaurally uncorrelated noise standard, providing a static interaural correlation cue. Subjects are able to perceive this increased envelope correlation as a narrowing of the spatial image in the high-frequency channels. In contrast to the static interaural envelope correlation cue, the long-term interaural (fine structure) correlation of the Phasewarp at the output of auditory filters is 0. The short-term, running interaural correlation of the Phasewarp oscillates with the modulation frequency as shown in Figure 1.

To estimate the effect of interaural envelope correlation for human psychophysics, the output of a 2000 Hz fourth-order gammatone filter with an ERB of 240.6 Hz (Patterson, 1994) was analyzed for different modulation frequencies. Because the effect solely depends on the modulation frequency/ERB ratio, and because the transfer functions of gammatone filters scale with their ERB, the analysis of a single exemplary filter is sufficient to predict the interaural envelope correlation as a function of modulation frequency in all auditory filters. The interaural envelope correlation of the Phasewarp was analyzed for ratios of modulation frequency/ERB of [0.01 0.02 0.04 0.08 0.16 0.32 0.64 0.96 1.28 1.96 2.56 3.84 5.12 10.24]. The standard for the control experiment, the PWEC noise, was then generated by mixing a diotic and an interaurally uncorrelated noise with a frequency dependent mixing ratio in the Fourier domain. The mixing ratio, mix, of the frequency component, freq, was as follows: mix = exp (−1.1 × modulation frequency/ERB), with ERB = 24.7 + freq/9.265 (Moore et al. 1990). The mixing function was found empirically and was adjusted to minimize the differences in the interaural envelope correlation of the PWEC noise and the Phasewarp stimulus at the output of the 2000 Hz gammatone filter for the above-mentioned modulation frequency/ERB ratios. The root-mean-square (RMS) difference of the envelope correlation of the Phasewarp and the PWEC noise

was minimized to 1.6%. As a result, the PWEC noise has the same envelope correlation in all filters of a gammatone filterbank as the Phasewarp, but it lacks the binaural modulation of the fine-structure correlation along the time and ITD axis of the Phasewarp.

Psychophysics. In an adaptive three-interval, three-alternative forced-choice procedure, four listeners (21–24 years in age) were asked to detect the interaurally modulated and the monaurally modulated stimuli following a two-down, one-up rule. This procedure estimates the 70.7% correct point on the psychometric function. The dependent variable was the cross-fading ratio of the interaurally or monaurally modulated stimuli and an interaurally uncorrelated noise with the same RMS values. The cross-faded stimulus $s_{fade}(t)$ was generated following the equation: $s_{fade}(t) = n(t) \times (1 - f) + s(t) \times f$, where $n(t)$ is the interaurally uncorrelated noise and $s(t)$ is the modulated stimulus. The cross-fading ratio, f, was adjusted in terms of dB [$20 \log_{10}(f)$], using an adaptive step size of 4, 2, or 1 dB. A cross-fading ratio of 0 dB ($f = 1$) refers to only the modulated stimulus, whereas at −6 dB ($f = 0.5$) the modulated and interaurally uncorrelated stimuli are added at a 1:1 ratio. The cross-fading ratios were later converted to signal-to-masker (SMR) ratios for data analysis. A cross-fading ratio of −6 dB corresponds to a 0 dB SMR. Preliminary experiments have revealed that the perceptual distances between different cross-fading ratios in decibels were more or less equal, in contrast to perceptual distances between decibel steps when a direct variation of the SMR was used. However, because the data presentation seemed more intuitive when plotted as SMR, the cross-fading ratios that were appropriate for the experiment were converted to SMR ratios. The experiment was performed with broadband versions of the modulated stimuli and with high- and low-pass-filtered versions of the Oscor and the Phasewarp. The cutoff frequencies for the low-pass-filtered stimuli was 1500 Hz, and the cutoff for the high-pass stimuli was 5000 Hz. Filters were implemented as "brick-wall" filters that, because of a Fourier transform-based algorithm, produce filter slopes in excess of 100 dB per octave. The remaining frequencies were filled with interaurally uncorrelated noise of the same spectrum level to preclude the use of off-frequency cues.

Stimulus duration was 1 s including 20 ms raised-cosine ramps. Listeners were seated in a sound-attenuated room. The stimuli were digitally generated on a personal computer using the AFC package for Matlab [developed at the Universität Oldenburg (Oldenburg, Germany) and the Technical University of Denmark (Copenhagen, Denmark)] at a sampling rate of 48 kHz and were presented from an RME Audio (Haimhausen, Germany) Digi 96/8 soundcard through Sennheiser (Wedemark, Germany) HD580 headphones at a sound pressure level of 60 dB sound pressure level (SPL). The headphones were digitally compensated to produce a frequency-independent response on a Bruel & Kjaer (Naerum, Denmark) 4153 artificial ear.

For each listener and stimulus condition, at least three adaptive runs were acquired. Individual data were only used for additional analysis if a threshold could be determined in all three runs. Data averaged across listeners are shown only if an average threshold could be determined for each tested listener.

Neurophysiology. Auditory responses from 87 single neurons in the DNLL were recorded from 26 adult Mongolian gerbils, *Meriones unguiculatus*. The detailed methods in terms of surgery, acoustic close-field stimulation, stimulus calibration, and recording techniques have been described previously (Siveke et al., 2006). All experiments were approved according to the German Tierschutzgesetz (AZ 55.2-1-54-2531-57-05).

The anesthetized animals (20% ketamine and 2% xylacine) were placed in a sound-attenuated chamber and mounted in a custom-made stereotaxic instrument (Schuller et al., 1986). Using motorized micromanipulators (Digimatic; Mitutoyo, Neuss, Germany) and a piezodrive (Inchworm controller 8200; EXFO Burleigh Products Group, Quebec,

Quebec, Canada) the electrode penetrations (tilted 10° or 5° laterally) were performed 1.3–2.6 mm lateral to the midline and 0.5–0.8 mm caudal of the interaural axis. Single-unit responses were recorded extracellularly using glass electrodes filled with 1 M NaCl (~10 MΩ). The amplified (7607; Toellner, Herdecke, Germany) and filtered (VBF/3; Kemo, Beckenham, UK) action potentials were fed into the computer via an A/D-converted (RP2.1; Tucker-Davis Technologies, Alachua, FL). Clear isolation of action potentials from a single neuron (signal-to-noise ratio, >5; see waveform of the recorded spikes in Figs. 4 and 8) was guaranteed by visual inspection (stable size and shape) on a spike-triggered oscilloscope and by off-line spike cluster analysis (Brainware; Jan Schnupp, Tucker-Davis Technologies).

Stimuli were generated at 50 kHz sampling rate by Tucker-Davis Technologies System III. Digitally generated stimuli were converted to analog signals (RP2.1; Tucker-Davis Technologies), attenuated (PA5; Tucker-Davis Technologies), and delivered to the earphones (Stereo Dynamic Earphones; MDR-EX70LP; Sony, Tokyo, Japan).

As search stimulus, interaurally uncorrelated noise bursts were used. To determine the best frequency (BF), randomized pure tone stimulations (nine frequencies; step size, BF/5; 10 dB steps) were presented. Both, the search stimulus and the frequency versus level response area were delivered binaurally with equal intensity at the ears (200 ms duration plus 5 ms raise–fall time; 2 Hz repetition rate). Response were defined as sustained if the neurons responded over the entire duration of the stimulus and not exclusively during the first 50 ms. Randomized broadband and narrowband (±10% around BF) noise stimuli (200 ms duration; 2 Hz repetition rate; 1 ms raise–fall time) were presented binaurally to determine the threshold (rate-level functions, 5 dB steps) and the maximal ITD [noise-delay functions, 0.15 (BF)-cycle steps] of the neurons. A unit was considered ITD-sensitive if the noise-delay function was modulated by ≥50% (i.e., if the minimum discharge rate was less than one-half of the maximum rate). To determine the response to the monaural amplitude and the binaural modulations, the stimulus level was set to 20 dB above the threshold of the neuron and the correlated noise condition to the best ITD of the neuron. For the first part (see Figs. 4–7), neuronal responses of 87 neurons to 10 repetitions of the monaural and binaural modulation stimuli with 1 s duration, a repetition period of 1.5 s, and 20 ms squared-cosine ramps were obtained. For the second part (see Figs. 8–10), neuronal responses of 19 neurons to 10 repetitions of the Oscor and Phasewarp stimuli with, as in the psychophysical study, additional interaurally uncorrelated noise at a SMR of −19, −16, −13, −10, −7, −3, 2, 10, and ∞ dB were obtained. To reduce the recording time, the stimulus duration (500 ms) and repetition period (750 ms) were decreased compared with the first part of the experiments. No differences of the neuronal response were observed using the first set of stimulation or the second set with an infinite SMR.

The vector strength (VS) for each modulation frequency of the monaural and the binaural modulations was calculated according to Goldberg and Brown (1968) and determined as significant if the $p < 0.001$ criterion in the Rayleigh test was fulfilled (Batschelet, 1991). For the second set of data (see Figs. 8–10), receiver operation characteristics (ROC) analysis (Green and Swets, 1966; Britten et al., 1992) was used to generate a neurometric function.

The standard condition was the presentation of the binaural stimuli at a SMR of −20 dB; the signal conditions were all presentations at higher SMRs. The ROC analysis was performed individually for each cell. For the rate-based analysis, first the proportion of repetitions in which the spike rate in the standard and signal condition exceeded a threshold criterion of, for example, one spike per second was calculated. For each SMR, ROC curves were generated by plotting the signal proportion against the standard proportion as a function of the threshold criterion. The area under the ROC curve corresponds to the neural performance in percent correct in a two-alternative, forced-choice task. Neurometric functions are obtained by plotting the percent correct as a function of SMR.

For the timing-based ROC analysis, the procedure is different. First, a 10-bin, normalized period histogram is calculated across all 10 repetitions of the stimulus at an SMR of infinity as a reference (P). Then, a 10-bin normalized period histogram (Q) is calculated for each repetition of each SMR, and the following Kullback–Leibler divergence (KL):
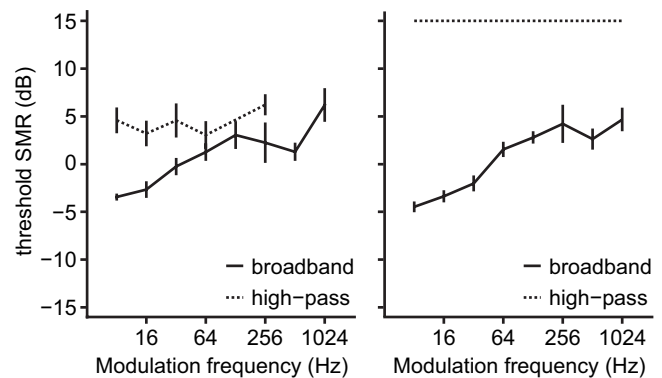

Figure 3. Control experiment concerning the contribution of perceptual binaural image width to binaural modulation detection. Results for the detection of the binaural modulation elicited by the Phasewarp stimulus in broadband and high-pass conditions are shown with interaurally uncorrelated noise in *A* and with the PWEC as standard stimuli in *B*. Although the listeners' performance in the broadband condition is unaffected by the different standard stimuli, none of the listeners could detect the Phasewarp in the high-pass condition with the PWEC as standard stimuli. Error bars indicate SEM.

$$\text{KL}(P,Q) = \sum_{x \in X} P(x) \times \log \frac{P(x)}{Q(x)}$$

is calculated between the reference period histogram and all other period histograms, where x corresponds to the 10 bins. If any bin, $P(x)$ or $Q(x)$ was 0, this value was replaced by 0.001. Thus, the KL divergence is typically small for a stimulus presentation at a high SMR and it increases with decreasing SMR. Then, the KL divergences calculated for the SMR of −20 dB were taken as the standard and the KL divergences for all higher SMRs were compared with this standard in the same way as it has been described in the previous paragraph for the rate-based analysis. It should be noted that the KL divergences between normalized period histograms are not completely independent of spike rate.

Because of the very long required stimulus duration and repetition rate and the high number of stimuli to be presented (144), only 10 repetitions per stimulus could be obtained. It should be noted that such a relatively low number of repetitions is not ideal for the ROC analysis.

Results

Psychophysics

The human sensitivity to a monaural modulation and three types of binaural modulation are shown in Figure 2A. Modulation detection was quantified in terms of the SMR at threshold with an interaurally uncorrelated noise masker. The data show that a monaural modulation (SAM) is detectable at a SMR of about −14 dB (Fig. 2A, dotted line). In line with previous findings (Viemeister, 1979), the threshold SMR increased for modulation frequencies >64 Hz. The three other plots depict human sensitivity to different types of binaural modulations (Fig. 2A, solid lines). Overall binaural sensitivity was worse than monaural sensitivity by ~10 dB SMR. Moreover, with the Oscor stimulus, sensitivity decreased relatively quickly with increasing modulation frequency >8–16 Hz. For modulation frequencies >128 Hz, the modulation in the Oscor stimulus (black line) could not be detected by all four listeners. This faster decrease in sensitivity compared with monaural sensitivity could be attributed to an additional binaural sluggishness with a time constant ≥10 ms, assuming a cutoff frequency <16 Hz in the data. Thresholds for binaural modulation detection with the Phasewarp stimulus are shown in light gray. For low modulation frequencies, thresholds were very similar to those with the Oscor, but for higher modulation frequencies, thresholds with the Phasewarp decreased only
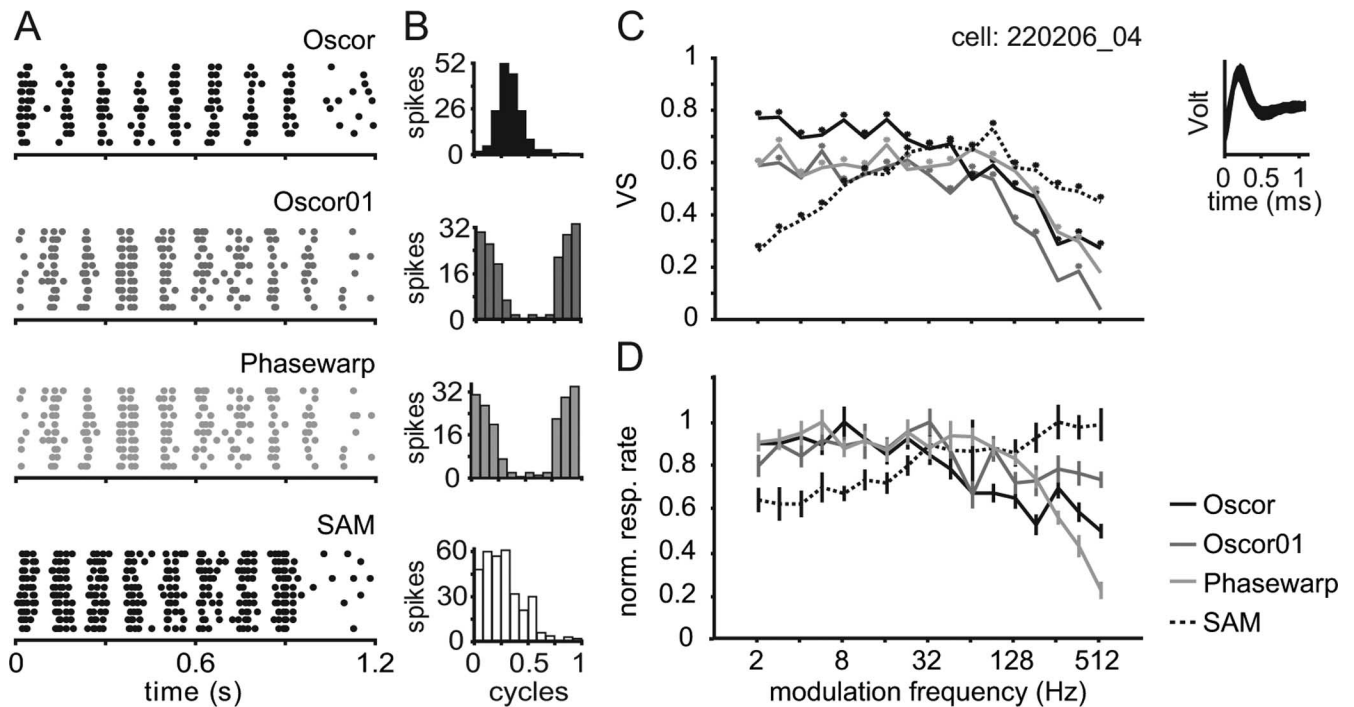
Figure 4. Single-cell responses of a DNLL neuron (BF, 1000 Hz) to the three types of binaural modulations and the monaural modulation elicited by SAM noise. Raster plots in response to stimuli with a modulation frequency of 8 Hz are shown in **A**, and period histograms are shown in **B**. **C** shows the temporal response characteristics, expressed as vector strength, and **D** shows the response rates of this neuron as a function of modulation frequency. The vector strength analysis reveals a low-pass characteristic with very similar sensitivity for the three types of binaural modulation and a bandpass characteristic for the monaural modulation. The response rate also shows a low-pass characteristic for the three types of binaural modulations but a slight high-pass characteristic for the SAM. The inset in **C** shows the highly stable spike waveform throughout the recording. Error bars indicate SEM. norm. resp. rate, Normalized response rate.

slowly and the modulations were detectable by all four listeners up to the highest modulation frequency of 1024 Hz.

The SMR at threshold can be converted to an amplitude modulation depth for the monaural amplitude modulation or an effective interaural modulation depth at threshold, and the inverted magnitude transfer function of a first-order low-pass filter can then be fitted to the transformed data. This fit, similar to Viemeister (1979), results in a time-constant estimate of ∼1.8 ms for the monaural amplitude modulation data. However, the low-pass characteristic of binaural modulation data elicited by the Phasewarp is considerably shallower than the 6 dB-per-octave slope of a first-order low-pass filter, which can be quantified with a time constant. Thus, no explicit time constant could be derived from the empirical data for the Phasewarp. However, the overall characteristic of the decrease in sensitivity with increasing modulation frequency is more similar to that of the monaural modulation than to that of the Oscor data. This shows that, using the Phasewarp stimulus, there was no indication of additional binaural sluggishness.

The fourth stimulus (dark gray) was a modified version of the Oscor stimulus denoted "Oscor01" (see Materials and Methods). The Oscor01 differed from the Oscor in that the modulation of interaural correlation only varies between 1 and 0 and not between 1 and −1. With the Oscor01, sensitivity to binaural modulation decreased dramatically. The binaural modulation was only detectable for the lowest presented modulation frequency of 8 Hz. A time constant or cutoff frequency cannot be estimated from this single data point. However, the data could be interpreted as indicating a considerably higher binaural sluggishness for the Oscor01 than found for the Oscor.

The psychophysical experiments were performed with broadband stimuli, and thus, it cannot be determined which frequency

range was used by the listeners to detect the binaural modulation. Therefore, the experiments were repeated with filtered versions of the Oscor and the Phasewarp. To avoid edge effects at the filter slopes and off-frequency listening, the filtered stimuli were added to inversely filtered, interaurally uncorrelated noise (see Materials and Methods). Results with the filtered stimuli are shown in Figure 2, *B* and *C*, for the Oscor and Phasewarp stimulus, respectively. In the low-pass condition (cutoff frequency, 1.5 kHz) (Fig. 2*B,C*, dashed lines), listeners' sensitivity to low modulation frequencies was unchanged compared with the broadband condition. However, the sensitivity to binaural modulations breaks down at lower modulation frequencies in the low-pass condition. This indicates that faster modulations were better preserved in frequency channels >1.5 kHz, or that the presence of the interaurally uncorrelated noise in the frequency region beyond 1.5 kHz lowered the performance.

All listeners could still discriminate the high-pass Phasewarp from the interaurally uncorrelated-noise standard for at least some modulation frequencies (Fig. 2*C*, dotted line). Some listeners could even detect the high-pass-filtered Phasewarp at all modulation frequency and some listeners could detect the high-pass-filtered Oscor at some modulation frequencies (data not shown). Because the human auditory system is no longer capable of encoding carrier (fine-structure) information at frequencies of 5 kHz and above (Moore, 1997), subjects must have been able to use a detection cue other than the interaural modulation. As described above (see Stimuli, Materials and Methods), the Phasewarp stimulus leads to an increasing interaural envelope correlation inversely related to the ratio modulation frequency/ERB with the ERB being approximately proportional to the center frequency of auditory filters. The decrease of the envelope correlation is perceivable by the auditory system at frequencies of
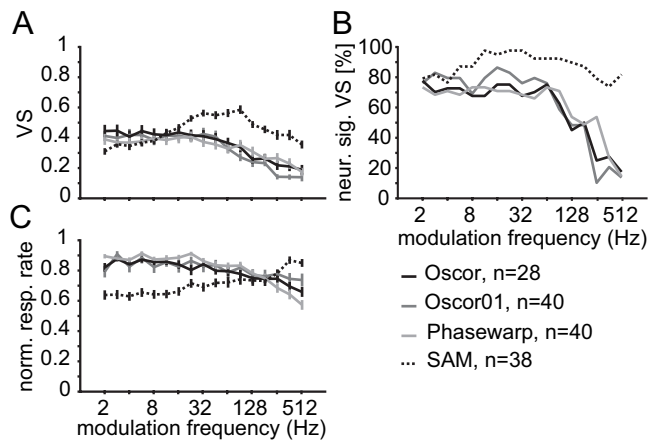
Figure 5. Population data of responses to the three types of binaural modulation and the monaural modulation. The average vector strength (**A**) and normalized response rate (**C**) are shown as a function of modulation frequency. **B** shows the percentage of recorded neurons with significant VSs for each of the four stimulus types as a function of modulation frequency. In the recorded population, most neurons could lock very well to the SAM modulation even at the highest modulation frequency (512 Hz) tested. For all three types of binaural modulations, the number of responses with significant VS similarly decreases with increasing modulation frequency for modulation frequencies above ~64 Hz. Error bars indicate SEM. norm. resp. rate, Normalized response rate; neur. sig. VS, percentage of neurons with significant VS.

5 kHz and higher. This decrease could have provided a static detection cue, a narrowing of the spatial image width, in comparison with the standard stimuli consisting of interaurally uncorrelated noise.

Results of a control experiment that precludes the (static) envelope correlation cue are shown in Figure 3. Three of the four listeners of the previous experiment participated. One additional subject who did not participate in the previous experiment was recruited. In this control experiment, the standard stimuli were PWEC noise, which was designed to elicit the same increase of interaural envelope correlation with increasing frequency as it is expected from the Phasewarp (see Materials and Methods). The data in the left panel show the performance with the standard stimulus used before (compare Fig. 2), and the data in the right panel show the performance with the PWEC noise as standard stimuli. The data with the PWEC noise show that, in the broadband condition, the high sensitivity to the binaural modulation elicited by the Phasewarp was preserved. In the high-pass condition with the PWEC-noise standard, however, none of the listeners could detect the binaural modulation anymore. This control experiment demonstrates that, although listeners could use the spatial image-width cue in the high-pass condition, the results of the main broadband experiment are not influenced by a potential usage of this cue.

In summary, the psychophysical data show that perceptual temporal resolution of binaural modulation is much better with the Phasewarp than with the classic Oscor stimuli and that the highest detectable rate of binaural modulation is comparable with that of monaural temporal processing.

Electrophysiology

Responses to monaural amplitude modulation and the three types of binaural modulation were obtained from a total of 87 single neurons in the gerbil DNLL. Only neurons exhibiting sustained responses, a best frequency <2 kHz, and ITD sensitivity to noise stimulations were investigated. The modulated stimuli were adjusted to the best ITD of each neuron as determined by
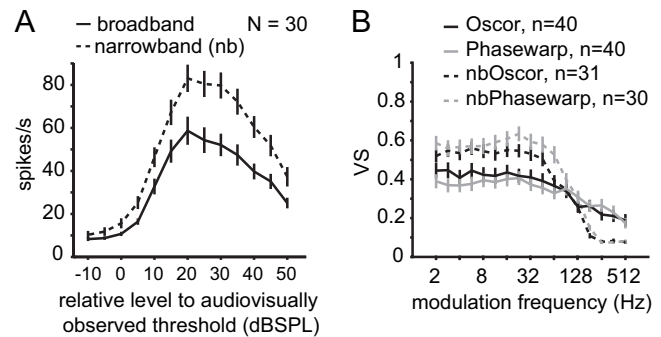

Figure 6. Effect of narrowband filtering on neural responses. The population rate-level-functions (**A**) show that, in general, neurons respond more vigorously to narrowband filtered stimuli than to broadband stimuli. The VS averaged across the population of neurons is shown as a function of modulation frequency in **B**. Whereas for low modulation frequencies, VS was generally higher in the narrowband than in the broadband condition, the narrowband filtering imposes a strong low-pass characteristic on the temporal response characteristics. This low-pass characteristic reflects a stimulus feature rather than a temporal limit of neural processing. Error bars indicate SEM.
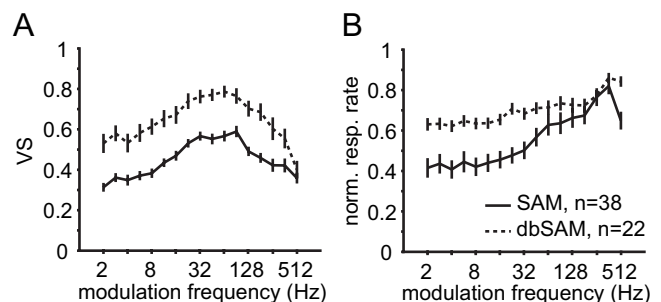

Figure 7. Effect of modulator shape on temporal and rate-response characteristics for monaural modulations. VSs are shown as a function of modulation frequency for two types of monaural modulations, namely a classical SAM and a sinusoidal amplitude modulation imposed on a logarithmic (decibel) amplitude scale. The population data show that the dBSAM produces more accurate temporal firing expressed as considerably higher VS. This temporal improvement is accompanied by a stronger increase in the rate response with increasing modulation frequency. Error bars indicate SEM. norm. resp. rate, Normalized response rate.

the noise-delay function. For low modulation frequencies, most neurons could synchronize their spike timing to the modulation period of either the SAM or the binaural modulation. Responses of a DNLL neuron are shown in Figure 4. The 8 Hz modulation of the stimuli is reflected both in the raster plots (Fig. 4A) and in the period histograms (Fig. 4B) of the neuronal responses. As a measure of how precise the neuronal response reflects the modulation of the stimulus, the VS was calculated from the period histograms. The VS as a function of modulation frequency is plotted in Figure 4C. For the three binaural modulations, the VS remained unchanged up to a modulation frequency of ~64 Hz and decreased with additional increasing modulation frequency. In contrast, for the SAMs, VS increased with increasing modulation frequency up to 64 Hz and decreased with additional increasing modulation frequency. Changes in the response rate as a function of modulation frequency are shown in Figure 4D. Whereas for the three binaural modulations, the response rate decreased slightly with increasing modulation frequency, the rate increased with increasing modulation frequency in response to SAMs.

Both the temporal and rate-response characteristics described for a single neuron are reflected in the population data shown in Figure 5. As in the single-neuron data, population VS as a func-
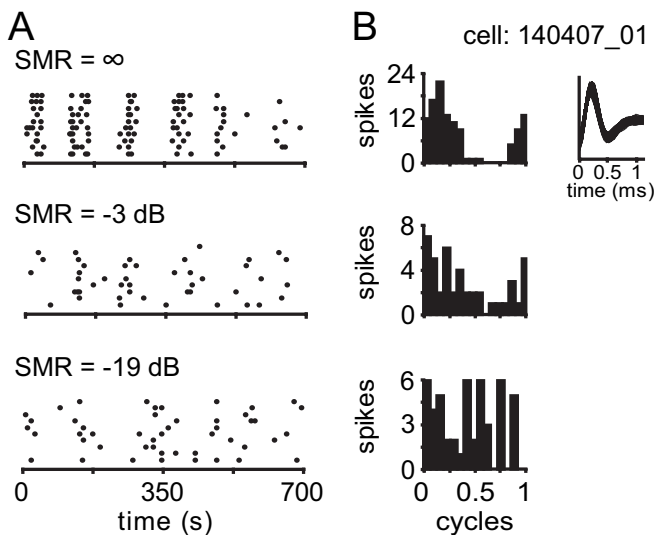
A

SMR = ∞

SMR = -3 dB

SMR = -19 dB

0   350   700
time (s)

B   cell: 140407_01

spikes

24
12
0

0 0.5 1
time (ms)

spikes

8
4
0

spikes

6
3
0

0   0.5   1
cycles

Figure 8. Responses of a DNLL neuron (BF, 800 Hz) to the Phasewarp stimuli at a modulation frequency of 8 Hz and different SMRs. Raster plots are shown in ***A*** and period histograms extracted from these responses are shown in ***B***. At an SMR of infinity, the spikes are well locked to binaural modulation, but this locking deteriorates with decreasing SMR. The inset in ***B*** shows the highly stable spike waveform.

tion of modulation frequency showed a low-pass characteristic with a cutoff frequency of ~32–64 Hz in response to binaural modulations (Fig. 5A). In response to SAMs, population VS had a bandpass characteristic with a best modulation frequency around 64 Hz. The percentage of neurons in each population with significant VSs is shown as a function of modulation frequency in Figure 5B. The majority of neurons (>60%) in each population exhibited significant VS for modulation frequencies up to 64 Hz. For higher modulation frequencies, the proportion of neurons locking to the modulation decreased rapidly for the binaural modulations and only slightly for SAMs. However, up to a modulation frequency of 512 Hz, still >15% of the neurons expressed significant VSs to the binaural modulations and 81% expressed significant VSs to the SAM. As it was the case for the single-neuron data in Figure 4, differences between the SAM and binaural modulations were visible in the population rate response (Fig. 5C): With increasing modulation frequency, the normalized response rate increased in response to SAM, but it decreased in response to binaural modulations.

DNLL neurons typically responded more vigorously to narrowband than to broadband stimulation. This is reflected in the population rate-level function shown in Figure 6A. Therefore, we tested whether the stimulus bandwidth affects the response to the binaural modulations, specifically, the Oscor and Phasewarp stimuli. In general, the VSs for the narrowband stimuli were higher than for the broadband stimuli (Fig. 6B). However, narrowband filtering imposes a low-pass characteristic in the modulation-frequency domain. This low-pass characteristic affected both monaural and binaural modulations in the same way (Fig. 6B). Thus, the steep decrease of the VSs to modulation frequencies >64 Hz does not reflect a low-pass characteristic of neural processing.

The VS for modulation frequencies <64 Hz was constant for the three binaural modulations in contrast to the monaural modulation. This may be attributable to the shape of the modulation, which is sinusoidal along a linear amplitude axis. As it is evident from the raster plot in Figure 4A, the responses to the SAMs had a higher "duty cycle" than the responses to the binaural modula-

tions. This high duty cycle results from the compressive characteristics of the auditory periphery. Peripheral compression can be circumvented by modulating the amplitude sinusoidally along the dB axis (dBSAM). Studying a subgroup of neurons ($n = 22$) with dBSAM stimuli, it was found that the VS strongly increased compared with SAM stimulation (Fig. 7A). Nevertheless, VS still increased with increasing modulation frequency up to 64 Hz. The lower duty cycle of the dBSAM stimulus in contrast to the SAM stimulus is reflected in the lower response rate of low modulation frequencies (Fig. 7B).

In summary, both the temporal and rate characteristics of the recorded neurons indicated a difference in sensitivity and precision of the encoding of monaural amplitude modulations compared with binaural modulations. Specifically, the analysis in Figure 5A showed that the VSs in response to SAMs at higher modulation frequencies were higher than in response to binaural modulations. The proportion of neurons that produced significant VS at high modulation frequencies in response to monaural modulations was higher than in response to binaural modulations (Fig. 5B). These results appear indicative of a neural correlate of binaural sluggishness. In contrast to the psychophysical data, all three types of binaural modulations elicited very similar responses in the DNLL.

Comparison of psychophysical and electrophysiological performance

The human psychophysical performance and the electrophysiological sensitivity of gerbil DNLL neurons are directly compared using a ROC approach. The ROC analysis has been successfully used in a number of studies relating physiology to psychophysics (Britten et al., 1992; Skottun et al., 2001; Firzlaff et al., 2006). The neurometric function reflects the probability of an ideal observer to accurately detect the modulation based on neuronal responses. To generate neurometric functions, additional electrophysiological data were recorded from 19 neurons for the Oscor and the Phasewarp stimuli at modulation frequencies and SMRs matching the range of psychophysical data acquisition. Raster plots and period histograms of a single-cell response with Phasewarp stimulation are shown in Figure 8. Both the precision of spike timing and the response strength increased with increasing SMR.

The ROC analysis was applied either on the strength or on the timing of the recorded responses (see Materials and Methods). Examples of neurometric functions generated from the data in Figure 8 and neurometric functions averaged across the population of 19 neurons are shown in Figure 9. Neurometric thresholds were extracted using the same threshold criterion (70.7% correct) as in the psychophysics (Fig. 9, dashed lines).

A direct comparison of electrophysiological and psychophysical thresholds is shown in Figure 10. Electrophysiological thresholds in the left column were based on a response-strength analysis, and electrophysiological thresholds in the right column were based on a response-timing analysis. Overall, electrophysiological thresholds are higher than psychophysical thresholds. However, both with the rate-based and with the timing-based analysis, differences are found between electrophysiological sensitivity to the binaural properties of the Oscor and Phasewarp. In qualitative agreement with the psychophysical data, electrophysiological sensitivity to the Phasewarp is higher and persists to higher modulation frequencies compared with the Oscor. With the spike-timing analysis, no thresholds could be determined for Oscor stimulation at a modulation frequency of 32 Hz because the neurometric function averaged across the cell population just failed to reach the 70.7% threshold criterion (compare Fig. 8F).
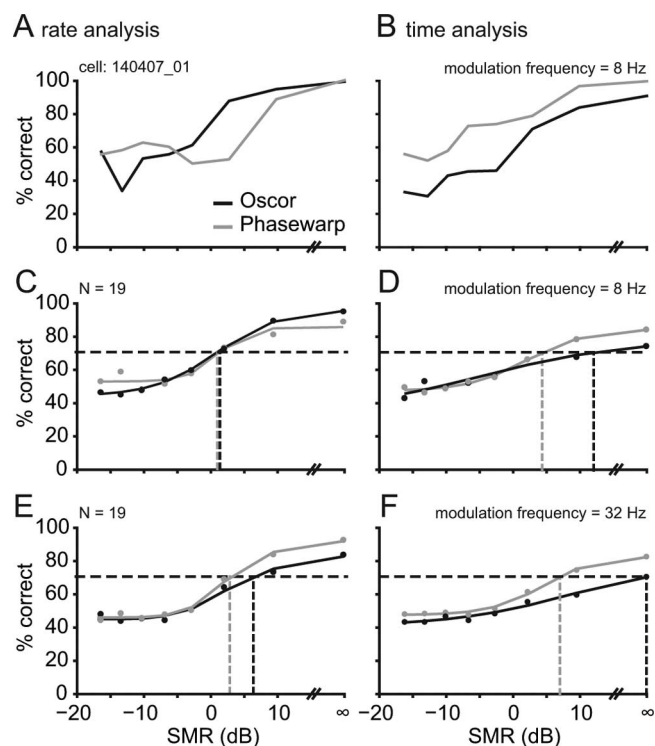
Figure 9. Neurometric functions based on a ROC analysis of the neuronal responses recorded as a function of SMR. Neurometric performance based on a rate analysis is shown in the left column; performance based on a temporal analysis is shown in the right column. Neurometric performances of a single cell (*A*, *B*; same neuron as shown in Fig. 8) and averaged across the population of neurons recorded (*C*, *D*) at a modulation frequency of 8 Hz expressed as percentage of correct choices of an ideal observer, are plotted as a function of SMR. Average neurometric functions in response to 32 Hz binaural modulations are shown in *E* and *F*. The horizontal dashed lines represent the threshold criterion (70.7% correct). The vertical dashed lines indicate the threshold SMRs extracted from a sigmoidal fit to the averaged psychometric functions. Note that, at a modulation frequency of 32 Hz, the averaged psychometric function based on the temporal analysis of the Oscor responses does not reach the threshold criterion (*F*).

Discussion

This study was designed to directly compare the temporal accuracy of the monaural and binaural system in following modulations of amplitude and interaural correlation, respectively. A combination of human psychophysics and gerbil electrophysiology was used. A set of three different binaural stimuli was created, none of which provided any monaural cues for the modulation.

Psychophysical sensitivity to the binaural modulations depended strongly on the stimulus type and the modulation frequency. Sensitivity was worst for the Oscor01 stimulus in which the modulation of interaural correlation was limited to the range between 0 and 1. An interaural correlation of 0 produces no focused binaural image, whereas a correlation of 1 produces a well focused, centralized binaural image. The psychophysical data show, however, that this modulation can only be detected at very low modulation frequencies (8 Hz). In the Oscor stimulus, the correlation is sinusoidally modulated between −1 and 1. A correlation of −1 produces a semifocused binaural image, typically at a different position in the head than the highly focused image produced by a correlation of +1. Thus, the Oscor produces a complex pattern of changes in spatial image width and position. The psychophysical sensitivity to the Oscor modulation was much better than to the Oscor01. Additionally, all listeners could detect modulations up to 128 Hz. The Phasewarp stimulus produces a time course of interaural correlation that is most similar

to a movement of a sound-source around the head. With the Phasewarp stimulus, sensitivity to binaural modulation for high modulation frequencies was considerably better than with the Oscor stimulus. Moreover, the low-pass characteristic of the sensitivity to the binaural modulation of the Phasewarp stimulus is most comparable with the low-pass characteristic of the sensitivity to the monaural modulation quantified with SAM noise. This finding suggests that, given the optimal stimulation, the binaural system is not more sluggish than the monaural system.

The current results appear to be in conflict with the results of previous investigations of binaural temporal processing. Grantham (1982) used Oscor stimuli to quantify the time constant of ITD extraction and found a considerably lower temporal resolution. However, his study was performed with stimuli that were first filtered and then modulated and added. The modulation, applied after the filtering, produces side bands that become increasingly audible with increasing modulation frequency. The author was well aware of this and confined his analysis to that range of modulation frequencies that were supposedly not contaminated with this spectral effect. Grantham concluded that the time constant of binaural processing was considerably longer than those described for monaural processing (Grantham and Wightman, 1979; Dau et al., 1999). He also attempted to reconcile his results with previous studies by Pollak (1978) who found that listeners could detect the periodic switching between binaural sound sources with periods as short as 1.5–2 ms. More recent quantifications using a masking-period pattern paradigm (Kollmeier and Gilkey, 1990) or a binaural probe configuration (Akeroyd and Summerfield, 1999; Akeroyd and Bernstein, 2001) suggested a considerably longer window of temporal integration for the binaural system than for the monaural system.

The apparent lack of binaural sluggishness in the current data does to some extent result from the difference in the experimental paradigm: In contrast to many previous studies, the present paradigm does not require listeners to make fine judgments in terms of, for example, minimum audible movement angle (Perrott and Musicant, 1977; Grantham, 1986) or localization blur (Blauert, 1972). Instead, it is a simple detection task, comparable with the detection of monaural amplitude modulation. The apparent discrepancy of our results with those by, for example, Grantham (1982) or Akeroyd and Summerfield (1999) may be related to the fact that these studies did not specifically implement time-varying ITDs, but rather periodic or abrupt changes in interaural correlation. It appears that the perceptual sensitivity to binaural modulation profits strongly from additional modulation along the ITD axis.

Without additional experiments, it is hard to speculate why pronounced sensitivity differences were found across the stimuli of the current study. To answer this question, it might, however, be helpful to inspect the right column of Figure 1. One hypothesis is that the sensitivity to binaural modulation depends on the prevalence of focused binaural images in the modulation cycle. The Oscor01 is modulated between a correlation of 0 and 1, and thus, during about one-half the modulation period, the stimulus produces no focused binaural image. The classic Oscor is modulated between correlation of −1 and 1, and thus, this stimulus produces one focused binaural image and an additional, semifocused image (correlation of −1) per modulation period. The Phasewarp, in contrast, creates a focused image at all times during the modulation period. Additionally, as can be seen from the right panel of Figure 1C, the correlation modulation occurs in a regular manner along the ITD axis.

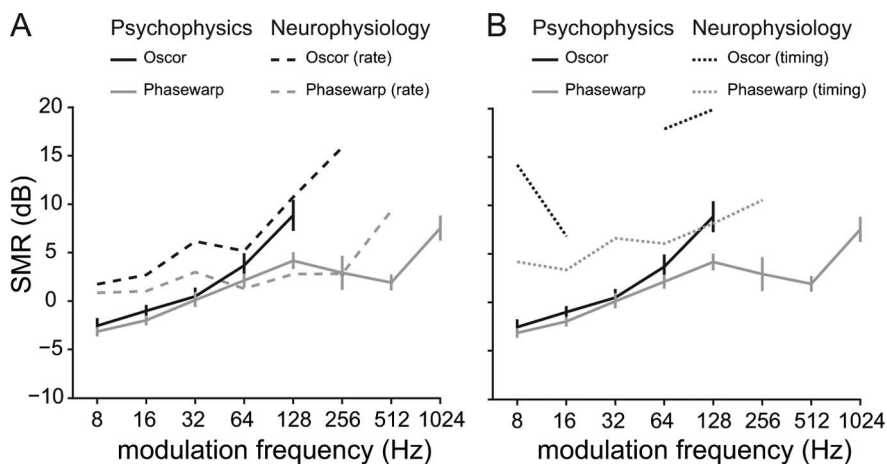The first set of the current electrophysiological results shows

Figure 10. Direct comparison of the human psychophysical performance and the gerbil neurometric performance for the detection of binaural modulations. Neurometric performance based on the rate analysis is shown in *A* and based on the temporal analysis in *B*. Although the overall neurometric performance is slightly worse than the human psychophysical performance, both rate and temporal analyses of neural data indicate better thresholds for the Phasewarp than for the Oscor stimuli. Thus, the neurometric data reflect the human psychophysical performance qualitatively correct.

that, at the level of the gerbil auditory midbrain, the monaural modulations of the SAM noise and the binaural modulations of the Oscor and Phasewarp stimuli are well preserved. This ability of midbrain neurons to follow fast temporal modulations has been shown previously for SAM stimuli (Langner and Schreiner, 1988; Krishna and Semple, 2000) and recently for the Oscor stimulus (Joris et al., 2006). It is noteworthy that, although the BFs of the recorded neurons were <2 kHz, many neurons showed significant VS to modulation frequencies as high as 512 Hz. The recordings with the bandpass filtered stimuli already show that filtering imposes a low-pass characteristic in the modulation domain. It appears that the bandpass filters in the gerbil inner ear are, despite the low BF, wide enough to transmit such high modulation frequencies. Averaged across the cells recorded, BF was 1 kHz; absolute threshold was at 28 dB SPL, and the Q20 bandwidth was 1.6 (i.e., the bandwidth was between 0.5 and one octave). Overall, the temporal code for monaural modulation was considerably better than for the binaural modulations. At higher modulation frequencies, both the VS averaged across the population and the number of units exhibiting significant VS are higher for SAM noise stimulation than for stimulation with any kind of binaural modulation. This difference could be interpreted as a neural correlate of binaural sluggishness.

In the recent electrophysiological study on potential sluggishness of the mammalian binaural system by Joris et al. (2006), single cells in the cat inferior colliculus were recorded while stimulating with the Oscor stimulus. Consistent with the present results, Joris et al. reported a very good temporal encoding of the binaural modulations. The discrepancy of these electrophysiological findings with the strong binaural sluggishness reported in the human-psychophysical work by Grantham (1982) was interpreted in that there is no neural substrate at the level of the midbrain or higher auditory stages to read out this temporal code. The current psychophysical data indicate, however, that the Phasewarp stimulus is capable of producing perceivable modulations of ITDs at modulation frequencies up to 1024 Hz. These new results indicate that, with optimal stimulation, the auditory system is well capable of reading out the temporal code of binaurally sensitive units when the time course of activation along both the ITD and the time axis is appropriate.

One crucial difference in psychophysical and electrophysiological approaches as used in the first set of electrophysiological data here or in Joris et al. (2006) is the way in which sensitivity is tested and expressed. In psychophysics, sensitivity is quantified as the SMR that is required to detect the modulation. In the electrophysiology, stimuli were always at an infinite SMR and the sensitivity is quantified as VS. Another difference is that psychophysics in humans is compared with electrophysiology in gerbils. The current working hypothesis is, however, that differences and similarities between monaural amplitude modulation coding and binaural modulation coding might persist in both mammalian species.

The second electrophysiological data set was recorded to quantitatively compare psychophysical and electrophysiological binaural modulation detection. Here, not only the modulation frequency but also the signal-to-noise ratio was varied in a manner consistent with the psychophysics. A ROC analysis based either on the response strength or on the spike timing of single neurons revealed significant differences between the encoding of the binaural modulations of Oscor and Phasewarp. The neurometric thresholds are in at least qualitative agreement with the psychophysical thresholds. Note that these differences are not visible when the stimuli are only presented at full modulation depth, as it is typically done in electrophysiological experiments. The ROC analysis based on the response strength provides a better fit to the psychophysical data than the analysis based on spike timing. Nevertheless, it may be critical to associate the psychophysical performance with the response strength: As outlined in the psychophysical control experiment, the degree of interaural correlation in higher frequency channels, in which the binaural system is mostly driven by the stimulus envelope, is lower for the Phasewarp than for interaurally uncorrelated noise. In consequence, the degree of interaural correlation at these frequencies decreases with decreasing SMR. Because binaural neurons can be driven by the stimulus envelope (Batra et al., 1989) and are sensitive to interaural correlation per se (Shackleton et al., 2005), they could pick up these changes.

Together, the current psychophysical and electrophysiological data show that, given appropriate binaural stimulation, the binaural system can process time-varying ITDs as fast as the monaural system can process time-varying amplitudes. The limits of processing were measured using detection thresholds. At binaural or monaural modulation frequencies >8 Hz, neither the monaural nor the binaural system is able to "track" the modulation as specific level fluctuation, binaural movement, or spatial image width change. Rather, a monaural or binaural flutter or roughness is perceived. If the term binaural sluggishness is used and related to time constants ≥10 Hz, the authors claim that this does not describe the temporal resolution of internal binaural information accessible to higher stages of the brain. More likely, an additional binaural sluggishness is task dependent and occurs if the task is to "track" specific interaural changes, as tracking a source in space or to perceive movement. The electrophysiological data show that neural encoding at the level of the DNLL is very fast. Our capability to read out these fast modulations, however, appears to depend strongly on the type of ITD change. The

readout appears to be much more effective when a focused binaural image is present at all times, whereas the readout appears sluggish when a binaural image emerges and disintegrates periodically. The ROC analysis and resulting neurometric thresholds show that both neural response strength and timing transmit information that substantiates the perceptual differences in the processing of binaural modulations.

References

Akeroyd MA, Bernstein LR (2001) The variation across time of sensitivity to interaural disparities: behavioral measurements and quantitative analyses. J Acoust Soc Am 110:2516–2526.

Akeroyd MA, Summerfield AQ (1999) A binaural analog of gap detection. J Acoust Soc Am 105:2807–2820.

Batra R, Kuwada S, Stanford TR (1989) Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. J Neurophysiol 61:257–268.

Batschelet E (1991) Circular statistics in biology. London: Academic.

Blauert J (1972) On the lag of lateralization caused by interaural time and intensity differences. Audiology 11:265–270.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765.

Dau T, Verhey J, Kohlrausch A (1999) Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. J Acoust Soc Am 106:2752–2760.

Ewert SD, Dau T (2000) Characterizing frequency selectivity for envelope fluctuations. J Acoust Soc Am 108:1181–1196.

Firzlaff U, Schornich S, Hoffmann S, Schuller G, Wiegrebe L (2006) A neural correlate of stochastic echo imaging. J Neurosci 26:785–791.

Goldberg JM, Brown PB (1968) Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study. J Neurophysiol 31:639–656.

Grantham DW (1982) Detectability of time-varying interaural correlation in narrow-band noise stimuli. J Acoust Soc Am 72:1178–1184.

Grantham DW (1986) Detection and discrimination of simulated motion of auditory targets in the horizontal plane. J Acoust Soc Am 79:1939–1949.

Grantham DW, Wightman FL (1978) Detectability of varying interaural temporal differences. J Acoust Soc Am 63:511–523.

Grantham DW, Wightman FL (1979) Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. J Acoust Soc Am 65:1509–1517.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Joris PX, van de Sande B, Recio-Spinoso A, van der Heijden M (2006) Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. J Neurosci 26:279–289.

Kohlrausch A, Fassel R, Dau T (2000) The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J Acoust Soc Am 108:723–734.

Kollmeier B, Gilkey RH (1990) Binaural forward and backward masking: evidence for sluggishness in binaural detection. J Acoust Soc Am 87:1709–1719.

Krishna BS, Semple MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. J Neurophysiol 84:255–273.

Kuwada S, Fitzpatrick DC, Batra R, Ostapoff EM (2006) Sensitivity to interaural time differences in the dorsal nucleus of the lateral lemniscus of the unanesthetized rabbit: comparison with other structures. J Neurophysiol 95:1309–1322.

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J Neurophysiol 60:1799–1822.

McAlpine D, Jiang D, Shackleton TM, Palmer AR (2000) Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation. J Neurophysiol 83:1356–1365.

Moore BC, Peters RW, Glasberg BR (1990) Auditory filter shapes at low center frequencies. J Acoust Soc Am 88:132–140.

Moore BCJ (1997) Introduction to the psychology of hearing. San Diego: Academic.

Palmer AR, Jiang D, McAlpine D, Shackleton TM (1998) Response to sound motion: interaural amplitude and phase modulation in the inferior colliculus. In: Psychophysical and physiological advances in hearing (Palmer AR, Rees A, Summerfield AQ, Meddis R, eds), pp 368–374. London: Whurr.

Patterson RD (1994) The sound of a sinusoid: spectral models. J Acoust Soc Am 96:1409–1418.

Perrott DR, Musicant AD (1977) Minimum auditory movement angle: binaural localization of moving sound sources. J Acoust Soc Am 62:1463–1466.

Pollak I (1978) Temporal switching between binaural information sources. J Acoust Soc Am 63:550–558.

Schuller G, Radtke-Schuller S, Betz M (1986) A stereotaxic method for small animals using experimentally determined reference profiles. J Neurosci Methods 18:339–350.

Seidl AH, Grothe B (2005) Development of sound localization mechanisms in the mongolian gerbil is shaped by early acoustic experience. J Neurophysiol 94:1028–1036.

Shackleton TM, Arnott RH, Palmer AR (2005) Sensitivity to interaural correlation of single neurons in the inferior colliculus of guinea pigs. J Assoc Res Otolaryngol 6:244–259.

Siveke I, Pecka M, Seidl AH, Baudoux S, Grothe B (2006) Binaural response properties of low-frequency neurons in the gerbil dorsal nucleus of the lateral lemniscus. J Neurophysiol 96:1425–1440.

Skottun BC, Shackleton TM, Arnott RH, Palmer AR (2001) The ability of inferior colliculus neurons to signal differences in interaural delay. Proc Natl Acad Sci USA 98:14050–14054.

Spitzer MW, Semple MN (1991) Interaural phase coding in auditory midbrain: influence of dynamic stimulus features. Science 254:721–724.

Spitzer MW, Semple MN (1993) Responses of inferior colliculus neurons to time-varying interaural phase disparity: effects of shifting the locus of virtual motion. J Neurophysiol 69:1245–1263.

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 66:1364–1380.