Articles in PresS. J Neurophysiol (April 20, 2005). doi:10.1152/jn.01226.2004

Title page

Type of article: Report

Microsecond precision of phase delay in the auditory system of the barn owl

Short title: Microsecond precision of phase delay

Hermann Wagner[1,2], Sandra Brill[1], Richard Kempter[3] and Catherine E. Carr[2]

[1] Institute for Biology II, RWTH Aachen, Aachen, Germany

[2] Department of Biology, University of Maryland, College Park, USA

[3] Institute for Theoretical Biology, HU Berlin, and Neuroscience Research Centre, Charité Medical Faculty of Berlin, and Bernstein Center for Computational Neuroscience, Berlin, Germany

This research was sponsored by the German Research Foundation (DFG, Wa-606/12, Ke-788/1-3), and by NIH DC00436 and NIH P30 04664 to CEC.

The auditory system encodes time with sub-millisecond accuracy. To shed new light on the basic mechanism underlying this precise temporal neuronal coding, we analyzed the neurophonic potential, a characteristic multiunit response, in the barn owl's nucleus laminaris. We report here that the relative time measure of phase delay is robust against changes in sound level, with a precision sharper than 20 microseconds. Absolute measures of delay, such as group delay or signal-front delay, had much greater temporal jitter, for example due to their strong dependence on sound level. Our findings support the hypothesis that phase delay underlies the sub-millisecond precision of the representation of interaural time difference needed for sound localization.

The barn owl is well known for its superb sound-localization capabilities [4, 6, 10, 13-17, 22, 23, 25]. The owl uses interaural time difference (ITD) to encode sound azimuth, with a behavioural accuracy of less than 10 μs [4] and a neuronal sensitivity of 25-100 μs [4, 16]. The highest overall monaural temporal sensitivity has been measured in the third-order nucleus laminaris (NL) where binaural convergence creates tuning to ITD [6, 18, 20, 23]. The NL exhibits a characteristic frequency-following multiunit response termed the neurophonic [20, 21, 23], which we used to study timing. The neurophonic well represents both monaural and binaural temporal sensitivity [23].

Earlier measurements of the conduction time from the ear to NL have found values between 2 and 3 ms [6]. Sullivan and Konishi [22] already mentioned the importance of phase, but did not quantify temporal precision. Koppl [15] quantified temporal precision in the second-order nucleus magnocellularis that provides input to the nucleus laminaris, but found that the response delay depended on sound level. The rate of change amounted to

about 5 μs/dB around the characteristic frequency. Taking into account that the range of interaural level differences experienced by barn owls is around 20 dB [13, 25], how is it possible that the ITD can be represented in NL of owls with a neuronal precision much sharper than 100 μs?

The auditory system encodes delay both as conduction time, an absolute time measure [9, 11, 19], and by phase locking, a relative time cue [2, 6, 15, 18, 22, 23]. Absolute and relative time codes are also known in physics, where a distinction is made between group velocity and phase velocity, thus resulting in group and phase delays [2, 9, 11, 15, 19]. While group delay describes the latency of the envelope of a bandpass-filtered signal, phase delay refers to the times of occurrence of its peaks and troughs. The high-frequency limit of group delay, the signal-front delay, has also been used as a measure of delay [9, 19]. We analyzed data obtained from the NL of the barn owl to determine which measure of delay was suited for precise and level-independent representation of ITD.

Nine barn owls (*Tyto alba pratincola*) were used in this study. The procedures conformed to NIH guidelines for animal research and were approved by the animal care and use committee of the University of Maryland. In contrast to earlier studies, the analogue waveform of the neurophonic potential in or close to NL was recorded at a sampling period of 20.8 μs with commercial, Epoxylite coated tungsten electrodes (Frederick Haer Corporation, Brunswick, ME) with impedances of 2-8 MΩ. Neurophonic recordings had the advantage of being stable for at least an hour and allowed measurements of local multiunit activity. Specific recording sites were defined by combining stereotaxic techniques, physiological characterization and histologically verified lesions.

Acoustic stimuli (clicks and noises) were digitally generated by custom-written software ("Xdphys" written in Dr. M. Konishi's lab at the California Institute of Technology, Pasadena, CA) driving a signal-processing system (Tucker Davies Technology, Gainesville, FL). Clicks had a rectangular form of varying intensity (0 dB (corresponding to 65 dB SPL) to 40 dB attenuation) and a duration of 2 samples (equivalent to 41.6 μs). Only condensation clicks were used. The standard click had 0 dB attenuation.

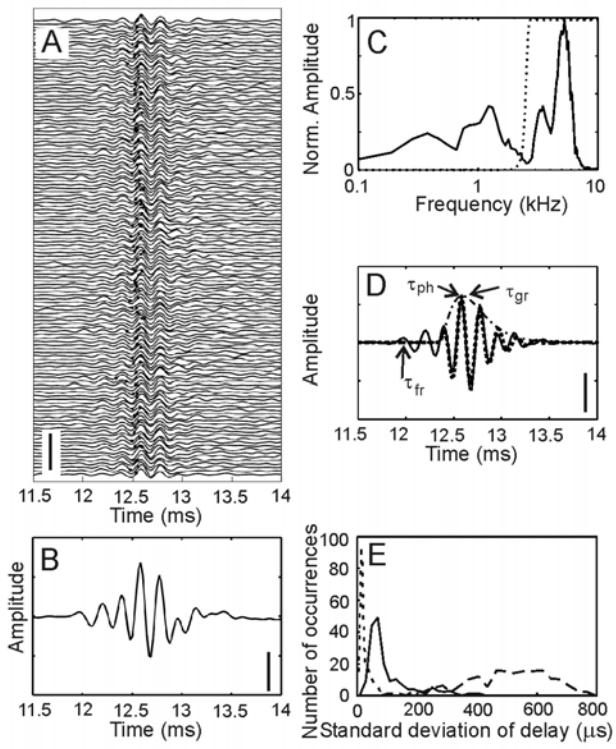

Figure 1. Analysis of neurophonic potentials evoked with standard clicks. Data from monaural stimulation. A) Time course of neurophonic potentials. The x-axis shows the time after recording onset. The click was presented at 10 ms after recording onset. The 128 individual traces are stacked on top of each other. Scale bar: 25 mV. B) Mean of the 128 traces shown in A. Scale bar: 2.5 mV. C) Normalized amplitude of the Fourier transform (512 points) of the mean response from 9.984 ms (data point 480) to the end of the recording. Note the different spectral components. The dotted line indicates the high-pass filter used for the separation of the components. D) High-pass filtered click-evoked mean response (solid line) and fit of this response with a Gammatone function of order 3 (dotted line). The envelope of the Gammatone function is shown by the dot-dashed line. The group delay (τ_{gr}) corresponds to the peak of the envelope. The signal-front delay (τ_{fr}) and the phase delay (τ_{ph}) were determined from the filtered, but unfitted data (see also text). Scale bar: 2.5 mV. E) Histogram of the variation of the three delay measures obtained from 96 recording sites (176 data sets, 89 left and 87 right stimulation, 128 repetitions each). Dotted line: phase delay, dashed line: signal-front delay, solid line: group delay. The bin width was 10 μs for the phase delay, 20 μs for the group delay, and 30 μs for the signal-front delay. Note that standard deviations are plotted, allowing for an analysis with a bin width that is smaller than the sampling interval.

Neurophonic responses to clicks were recorded in the 3.5-7 kHz region of the tonotopically organized NL. The spontaneous activity (10 ms before click presentation) as well as the driven activity (10 ms after click presentation) were stored. Clicks were repeated 128 times (Fig. 1A). The driven activity contained an oscillatory response (Fig. 1A, B). Its envelope increased smoothly within about 1 ms and fell off almost symmetrically. The oscillation under the envelope typically exhibited a complex waveform containing several spectral components. Fourier analysis showed that one or two components were below 2 kHz (Fig. 1C). Another component was close to the best frequency as obtained from iso-level frequency response curves. Since we wanted to study processes related to frequency tuning, only the high-frequency component was analyzed. Therefore, the neurophonic potential was high-pass filtered to reveal the oscillation of the high-frequency component alone (Fig. 1D). Auditory filtering is well described by gammatone functions and their derivatives [12, 24]. Thus, the high-pass filtered click-evoked response was fitted with a Gammatone function of order 3 (Fig. 1D).

In the click-evoked response typically several peaks and troughs could be distinguished. We considered only local extrema occurring after stimulus presentation and having an amplitude greater than the mean plus 2 standard deviations of the background noise for at least 3 consecutive data points. The latency of the first extremum detected in this way was the signal-front delay (Fig. 1D). The group delay was assigned to the maximum of the envelope of the Gammatone function (Fig. 1D). The extremum closest to the group delay, determined in the first of the 128 trials that was within one standard deviation of the average group delay, was chosen for the phase delay (Fig. 1D). To estimate the variability or "jitter" of each type of delay, we used the respective extrema in each of

the 128 traces (Fig. 1A). The standard deviation of these delays determined over the 128 repetitions was taken as a measure for the variability and was plotted as a data point in a histogram (Fig. 1E). In our sample of 176 data sets, the signal-front delay had the largest jitter (median: 500 μs), while the jitter in the phase delay was smallest (median: 10.4 μs). The jitter in the group delay (median: 64 μs) was between these two extremes. Thus, the temporal precision necessary for ITD coding [4, 16] could be achieved with the phase delay and the group delay, but not with the signal-front delay. Therefore, we did not consider the signal-front delay further.

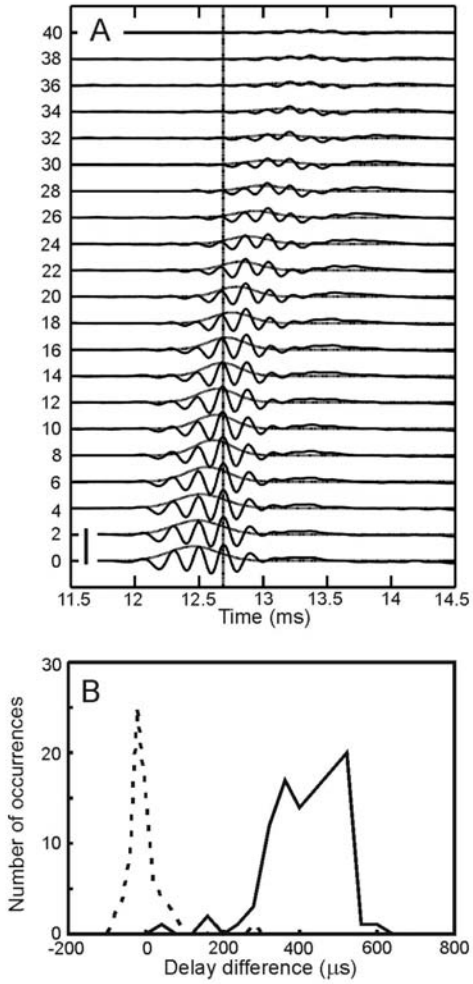

Figure 2. Level-dependence of click-evoked responses. Data from monaural stimulation. A) When the stimulus level decreased (numbers on left side refer to attenuation in dB), the shape of the waveform remained oscillatory, the amplitude decreased and the response occurred at longer latencies. However, as indicated by the vertical line, the position of the response maxima remained stable. Scale bar: 10 mV. B) Level dependence of delays in the pooled data of 72 data sets. Stimulus amplitudes were decreased from 0 dB to 20 dB attenuation. To determine the changes in phase or group delay induced by changes in click level, the times of arrival of the respective response extrema evoked by the different click intensities were picked from the mean curves and subtracted from each other. For phase delay values this sometimes included jumps between peaks as explained in the text. Each difference provided a data point for the histograms. Solid line: Group delay, dotted line: phase delay. Bin widths: 20 μs for the phase delay and 40 μs for the group delay.

In a second experiment, we decreased click amplitudes up to 40 dB. As is typically observed in audition, the group delay increased as the stimulus level decreased (Fig. 2). In the example shown in Fig. 2A, however, peak number 4 at 0 dB attenuation coincided with peak number 3 at 20 dB attenuation and with the barely visible peak number 1 at 40 dB attenuation (vertical line in Fig. 2A). Note that the definition of phase delay allows for a jumping between subsequent peaks [9]. In 72 data sets obtained from 43 recording sites phase delay remained essentially constant with a narrow distribution around one sample point (mean: -3 μs, Fig. 2B). In contrast, group delay increased up to 0.6 ms when click level was reduced by 20 dB (Fig. 2B), in agreement with the result in [6].

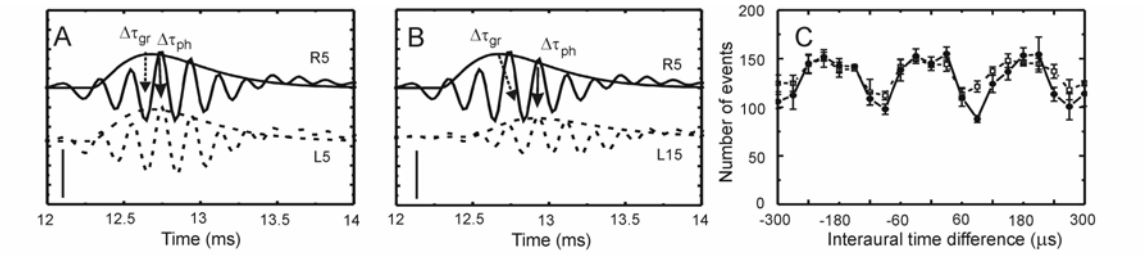

Figure 3. Influence of interaural level difference (ILD) on interaural time difference (ITD). Data from monaural stimulation (A, B) and binaural stimulation (C). A) At equal levels of the clicks presented to the left and right ears (5 dB attenuation: L5, R5), the difference between right and left group delays ($\Delta\tau_{gr}$) and phase delays ($\Delta\tau_{ph}$) were near zero microseconds. B) When the level on the left side was decreased by 10 dB (L15), group delay increased, while the phase delay stayed constant. C) In the binaural situation with broadband noise as the stimulus, the positions of the extrema of the ITD curve for an ILD of 0 dB (solid line) were almost identical to the extrema of the ITD curve measured with an ILD of 10 dB (dotted line). Events were obtained from an arbitrary setting of a TTL level trigger. Since $\Delta\tau_{gr}$ had changed, but the ITD curve did not change, the phase delay, but not the group delay is important for representing ITD. Scale bars in A and B: 2.5 mV.

The low variability and the level invariance of phase delay implied that this delay measure would be the most reliable code for representing the behaviorally relevant interaural time difference (ITD). The "best ITD" at a given recording site may be computed from two phase delays obtained through monaural stimulation [23]: we subtracted the phase delay for stimulation of the left ear from the phase delay for stimulation of the right ear.

This subtraction lead to a "best ITD" that was independent of interaural level difference (ILD) (compare Fig. 3A for 0 dB ILD with Fig. 3B for 10 dB ILD). On the other hand, ITDs between two group delays depended on the ILD. Even though an ITD computed from the group delay may be zero for 0 dB ILD (Fig. 3A), it changed significantly for 10 dB ILD (Fig. 3B).

Because the delay differences obtained from monaural responses in Figs. 3A and B only indirectly represented the binaural situation, we also tested the level-dependence of ITD tuning with binaural stimuli in a few (n=6) cases. In Fig. 3C we plotted the response at a given recording site as a function of the ITD of binaural noise, in contrast to our previous results obtained through neurophonic potentials in response to monaurally presented clicks. The ITD tuning curves for ILDs of 0 and 10 dB were virtually identical, in agreement with previous findings that ITD-tuning for binaural stimulation is stable under varying conditions of stimulus level [17, 25]. Thus, the level tolerance of ITD-tuning as shown in Fig. 3C independently demonstrates the importance of phase delay.

Thus, phase delay, and not group delay or signal-front delay, appears to underlie ITD tuning. Single unit recordings from auditory nerve and cochlear nucleus support this conclusion [15, 22]. Likewise, auditory nerve fibres of squirrel monkeys show very low phase-delay jitter when stimulated with sinusoids near the characteristic frequency [2].

The level independence of phase delay is consistent with a variety of filters proposed for peripheral auditory processing [12, 24]. The remarkable stability of phase delay is consistent with the model of Gerstner et al. [10] and Kempter et al. [14], who predicted that during development synapses from NM to NL and axonal arbors from NM to NL are selected in such a way that phase delays are similar. These authors also argued that

only such a selection allows for using phase delay to code temporal information in the NL and to represent ITD. Note that this conclusion is in line with the existence of a neurophonic potential in adult animals when it is assumed that the neurophonic potential is typically the summed response of an ensemble of magnocellular axons. A coherent summation of responses of different axons is only feasible when we have a coincident arrival of volleys of phase-locked spikes at the borders of NL and a coherent transmission of spikes through the nucleus. In other words, theory predicts that phase delays in different magnocellular axons must be similar.

Timing is important in many neuronal systems. It plays a role in models of learning (spike-timing dependent plasticity), in feature binding as well as in precise reactions to dynamic stimuli like approaching targets. To compare the precision of the different systems, a temporal quality factor is helpful. The coefficient of variation (CV), defined as the quotient of standard deviation and mean, may be a good criterion. In our example, the CV for the phase delay is around 0.01. In systems such as the visual cortex temporal jitter is much larger [3, 5] (CV~0.1), while values similar to the CV observed in the owl's NL are found in the electrosensory system [7], the auditory system of bats [8], and may be computed from synfire chains [1].

Acknowledgements

Michael Knepper and Ed Smith helped with measuring and analyzing click responses. Roland Schätte made helpful comments on the manuscript.

References

1. Abeles M, Bergman H, Margalit E, Vaadia E. Spatiotemporal firing pattern in the frontal cortex of behaving monkeys. J Neurophysiol 1993; 70: 1629-1638.

2. Anderson DJ, Rose JE, Hind JE, Brugge JF. Temporal position of discharges in single auditory nerve fibres within the cycle of a sine-wave stimulus: Frequency and intensity effects. J Acoust Soc Am 1971; 49: 1131-1139.

3. Bair W, Cavanaugh JR, Smith MA, Movshon JA. The timing of response onset and offset in macaque visual neurons. J Neurosci 2002; 22: 3189-3205.

4. Bala AD, Spitzer MW, Takahashi TT. Prediction of auditory spatial acuity from neural images on the owl's auditory space map. Nature 2003; 424: 771-774.

5. Bisley JW, Krishna BS, Goldberg ME. A rapid and precise on-response in posterior parietal cortex. J Neurosci 2004; 24: 1833-1838.

6. Carr CE, Konishi M. A circuit for detection of interaural time differences in the brainstem of the barn owl. J Neurosci 1990; 10: 3227-3246.

7. Carr CE, Heiligenberg W, Rose GTJ. A time-comparison circuit in the electric fish midbrain. I. Behavior and physiology. J Neurosci 1986; 6: 107-119.

8. Covey E, Casseday JH. The monaural nuclei of the lateral lemniscus in an echolocating bat: parallel pathways for analyzing temporal features of sound. J Neurosci 1991; 11: 3456-3470.

9. Fitzgerald JV, Burkitt AN, Clark GM, Paolini AG. Delay analysis in the auditory brainstem of the rat: comparison with click latency. Hearing Res 2001; 159: 85-100.

10. Gerstner W, Kempter R, van Hemmen JL, Wagner H. A neuronal learning rule for submillisecond temporal coding. Nature 1996; 383: 76-78.

11. Goldstein JL, Bear T, Kiang NYS. A theoretical treatment of latency, group delay, and tuning characteristics for auditory nerve responses to clicks and tones. In Physiology of the auditory system. Sachs MB (editor). National Education Consultants 1971. pp. 133-141.

12. Irino T, Patterson RD. A compressive gammachirp auditory filter for both physiological and psychophysical data. J Acoust Soc Am 2001; 109: 2008-2022.

13. Keller CH, Hartung K, Takahashi TT. Head-related transfer functions of the barn owl: measurement and neural responses. Hearing Res 1998; 118: 13-34.

14. Kempter R, Leibold C, Wagner H, van Hemmen JL. Formation of temporal-feature maps by axonal propagation of synaptic learning. Proc Natl Acad Sci USA 2001; 98: 4166-4171.

15. Koppl C. Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba. J Neurosci 1997; 17: 3312-3321.

16. Moiseff A, Konishi M. Neuronal and behavioral sensitivity to binaural time differences in the owl. J Neurosci 1981; 1: 41-49.

17. Pena JL, Viete S, Albeck Y, Konishi M. Tolerance to sound intensity of binaural coincidence detection in the nucleus laminaris of the owl. J Neurosci 1996; 16: 7046-7054.

18. Reyes AD, Rubel EW, Spain WJ. In vitro analysis of optimal stimuli for phase-locking and time-delayed modulations of firing in avian nucleus laminaris neurons. J Neurosci 1996; 16: 993-1007.

19. Ruggero MA. Systematic errors in indirect estimates of basilar membrane travel times. J Acoust Soc Am 1980; 67: 707-710.

20. Schwarz DWF. Can central neurons reproduce sound waveforms? An analysis of the neurophonic potential in the laminar nucleus of the chicken. J Otolaryng 1992; 21: 30-38.

21. Snyder RL, Schreiner CE. The auditory neurophonics: basic properties. Hearing Res 1984; 15: 261-280.

22. Sullivan WE, Konishi M. Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl. J Neurosci 1984; 4: 1787-1799.

23. Sullivan WE, Konishi M. Neural map of interaural phase difference in the owl's brainstem. Proc Natl Acad Sci USA 1986; 83: 8400-8404.

24. Tan Q, Carney LH. A phenomenological model for the responses of auditory-nerve fibres. II. Nonlinear tuning with a frequency glide. J Acoust Soc Am 2003; 114: 2007-2020.

25. Viete S, Pena JL, Konishi M. Effects of interaural intensity difference on the processing of interaural time difference in the owl's nucleus laminaris. J Neurosci 1997; 17: 1815-1824.