Estimation of cochlear response times using lateralization of frequency-mismatched tones

Strelcyk, Olaf; Dau, Torsten

Published in:
Acoustical Society of America. Journal

Link to article, DOI:
10.1121/1.3192220

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Strelcyk, O., & Dau, T. (2009). Estimation of cochlear response times using lateralization of frequency-mismatched tones. *Acoustical Society of America. Journal*, *126*(3), 1302-1311.
https://doi.org/10.1121/1.3192220

Estimation of cochlear response times using lateralization of frequency-mismatched tones

Olaf Strelcyk and Torsten Dau
Centre for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, Building 352, Ørsteds Plads, 2800 Kongens Lyngby, Denmark

(Received 3 April 2009; revised 1 July 2009; accepted 1 July 2009)

Behavioral and objective estimates of cochlear response times (CRTs) and traveling-wave (TW) velocity were compared for three normal-hearing listeners. Differences between frequency-specific CRTs were estimated via lateralization of pulsed tones that were interaurally mismatched in frequency, similar to a paradigm proposed by Zerlin [(1969). J. Acoust. Soc. Am. **46**, 1011–1015]. In addition, derived-band auditory brainstem responses were obtained as a function of derived-band center frequency. The latencies extracted from these responses served as objective estimates of CRTs. Estimates of TW velocity were calculated from the obtained CRTs. The correspondence between behavioral and objective estimates of CRT and TW velocity was examined. For frequencies up to 1.5 kHz, the behavioral method yielded reproducible results, which were consistent with the objective estimates. For higher frequencies, CRT differences could not be estimated with the behavioral method due to limitations of the lateralization paradigm. The method might be useful for studying the spatiotemporal cochlear response pattern in human listeners.

[DOI: 10.1121/1.3192220]

PACS number(s): 43.64.Ri, 43.64.Kc, 43.66.Pn, 43.66.Lj [BLM] Pages: 1302–1311

I. INTRODUCTION

The cochlea separates a sound into its constituent tonal components and distributes their responses spatially along its length by the distinctive spatial and temporal vibration patterns of its basilar membrane (BM). For example, the vibration pattern evoked by a single tone appears as a traveling wave (TW) (e.g., Ruggero, 1994; Robles and Ruggero, 2001). This wave propagates down the cochlea and reaches maximum amplitude at a particular point, before slowing down and decaying rapidly. The lower the frequency of the tone, the further its wave propagates down the cochlea. Hence, each point along the cochlea has a characteristic frequency (CF) to which it is most responsive. This tonotopic map is an important organizational principle of the primary auditory pathway and is preserved all the way to the auditory cortex (Clarey *et al.*, 1992).

At the level of the auditory nerve, the frequency of a tone is encoded both spatially, by its CF location, and temporally, by the periodicity of the responses in the nerve fibers that innervate the CF (cf. Ruggero, 1992). Several studies have suggested that the extraction of spatiotemporal information, i.e., the combination of phase-locked responses and systematic frequency-dependent delays along the cochlea (associated with the TW), may be important in the context of pitch perception (e.g., Loeb *et al.*, 1983; Shamma and Klein, 2000), loudness perception (Carney, 1994), localization (e.g., Shamma *et al.*, 1989; Joris *et al.*, 2006), speech formant extraction (e.g., Deng and Geisler, 1987), and tone-in-noise detection (e.g., Carney *et al.*, 2002). It has been proposed that a distorted spatiotemporal response might be, at least partly, responsible for the problems of hearing-impaired listeners to process temporal-fine-structure information (e.g., Moore, 1996; Moore and Skrodzka, 2002; Buss *et al.*, 2004). This may be one of the reasons for their difficulties to understand speech in noise. However, so far, empirical evidence for spatiotemporal information processing in humans is lacking since BM response patterns are difficult to monitor.

This study focused on one important component of the spatiotemporal BM response pattern: the cochlear response time (CRT) (e.g., Don *et al.*, 1993), which reflects the propagation delay of the TW. Consistent estimates of frequency-specific CRTs in humans have been obtained using different objective noninvasive methods, such as measurements of compound action potentials (e.g., Eggermont, 1976), stimulus-evoked otoacoustic emissions (e.g., Norton and Neely, 1987; Tognola *et al.*, 1997), tone-burst-evoked auditory brainstem responses (ABRs) (e.g., Gorga *et al.*, 1988), and derived-band click-evoked ABRs (e.g., Don and Eggermont, 1978; Parker and Thornton, 1978a; Eggermont and Don, 1980; Donaldson and Ruth, 1993; Don *et al.*, 1993).

Early psychoacoustic attempts to estimate CRTs or TW velocity were motivated by von Békésy's (1933) observation that the perceived position of clicks, presented to both ears, varied systematically when low-frequency masking tones were presented to one ear. Elaborating on this, Schubert and Elpern (1959) presented clicks in the presence of high-pass filtered noise with cutoff frequencies differing by half an octave between the two ears. The interaural time difference (ITD) that centered the unified percept at the midline was taken as an estimate of the difference in CRTs between the BM places corresponding to the noise cutoff frequencies in the two ears. However, the TW velocity derived from these CRT disparities was substantially larger than the TW velocity estimates obtained by means of the above mentioned objective methods (e.g., Donaldson and Ruth, 1993). As mentioned by Deatherage and Hirsh (1959) and Zerlin (1969), interaural loudness differences of the clicks might have influenced lateralization in the paradigms used by von Békésy (1933) and Schubert and Elpern (1959).

Instead of using click stimuli, von Békésy (1963b) and later Zerlin (1969) used pulsed tones that were interaurally mismatched in frequency. Both, von Békésy and Zerlin reported that listeners perceived the tones as fused, lateralized toward the ear receiving the higher-frequency tone. Zerlin measured the ITD needed to center the percept of the tones and took this as an estimate of the difference in CRTs between the BM places corresponding to the different tone frequencies in the two ears. The derived TW velocities were in good agreement with objective estimates of TW velocity (cf. Donaldson and Ruth, 1993). However, as noted by Neely *et al.* (1988), the reliability of Zerlin's estimates may be limited considering the difficulty of the psychoacoustic task and the fact that no further reports have been published since the original study in 1969.

If the lateralization of the interaurally mismatched tones reflected differences in CRTs, the paradigm would present a direct link between early cochlear disparities and spatial perception. Hence, particularly in view of the high temporal acuity of binaural auditory processing, which resolves ITD changes of less than 10 μs (Yost, 1974), this behavioral paradigm might serve as a complement to the objective measures of CRT mentioned above. Furthermore, Zerlin's (1969) paradigm bears a close relation to the concept of (across-ear) spatiotemporal processing. In both concepts, lateralization is supposed to be based on the comparison of information from mismatched frequency channels in the two ears. However, it is not clear if the lateralization in Zerlin's paradigm is based on interaural level differences (in the envelope at onset/offset), interaural time differences (in the fine structure), or a combination of both. Buus *et al.* (1984) suggested that temporal-fine-structure information during the first tone cycles might play a role in the lateralization of mismatched tones at low frequencies. This was supported by Magezi and Krumbholz (2008), who provided evidence that the binaural system can extract fine-structure information from interaurally mismatched frequency channels.

In the present study, behavioral estimates of CRT disparities and TW velocity were obtained for three normal-hearing listeners, using a similar paradigm to the one used by Zerlin (1969). In order to minimize measurement variability due to subjective listener criteria, an adaptive procedure was used to determine the ITD that centered the unified percept. The influences of loudness balancing, tone presentation level, and potential between-ear asymmetries on the CRT and TW velocity estimates were examined. For direct comparison, estimates of CRTs and TW velocities for the same listeners were obtained from derived-band ABRs. Since these estimates provide an objective "reference," they are presented first.

II. ABRs

A. Method

1. Listeners

The three female listeners were aged between 23 and 24 years and had audiometric thresholds better than 20 dB hearing level (ISO 389-8, 2004) at all octave frequencies from 125 to 8000 Hz and from 750 to 6000 Hz.

2. Stimuli

Rarefaction clicks were produced by applying 83-μs rectangular pulses (generated in MATLAB®) to an Etymotic Research ER-2 insert earphone. The clicks were presented monaurally at a level of 93-dB peak-to-peak equivalent sound pressure level (ppe SPL), with a repetition rate of 45 Hz. The acoustic clicks were calibrated using an occluded-ear simulator [IEC 60711, 1981; Brüel & Kjær (B&K) 4157] mounted with an ear-canal extension (B&K DP0370). Response latencies were corrected for a constant 1-ms delay introduced by the tubing of the ER-2 earphone.

Ipsilateral pink-noise masking was used to obtain derived-band ABRs (Don and Eggermont, 1978). High-pass noise maskers with cutoff frequencies of 0.5, 1, 2, 4, and 8 kHz were generated in the spectral domain as random-phase noise (with components outside the passband set to zero) and played back via a second ER-2 insert earphone, which was coupled to the first ER-2 earphone via an ER-10B+ transducer (without using the microphone). The spectrum level of the high-pass noise maskers was identical to that of the broadband pink noise, for which a level of 91 dB SPL was found to be sufficient to mask the ABR to the 93-dB ppe SPL clicks.

Perceptual click thresholds were measured for 500-ms click trains using a three-interval, three-alternative, forced-choice (3I-3AFC) task, tracking the 71%-correct point (one up, two down) on the psychometric function. The final threshold was estimated as the arithmetic mean over three runs. The average click threshold for the three listeners was 33.7 (31.5, 36.4) dB ppe SPL, with the values in parentheses representing the range of the individual results. These thresholds are lower than the corresponding reference threshold of 43.2 dB ppe SPL given by Richter and Fedtke (2005), which can be attributed to differences in click repetition rate and the different ear tips used. The ER1-14A used by Richter and Fedtke and the ER10-14 used in the present study differ in the diameter of the ear-tip tubes.

3. ABR recordings

Listeners lay on a couch in an acoustically and electrically shielded booth. The ABRs were measured differentially between electrodes applied to the vertex (C_z in the 10/20 system) and the ipsilateral mastoid (M_1 or M_2). Another electrode applied to the forehead (F_{pz}) served as ground. The electrode signals were acquired using a Neuroscan SynAmps 2 system, at a sampling rate of 20 kHz. Offline bandpass-filtering between 0.1 and 2 kHz (forward-backward filtering) was applied. Weighted averaging, as discussed in Elberling and Wahlgreen (1985) and in Don and

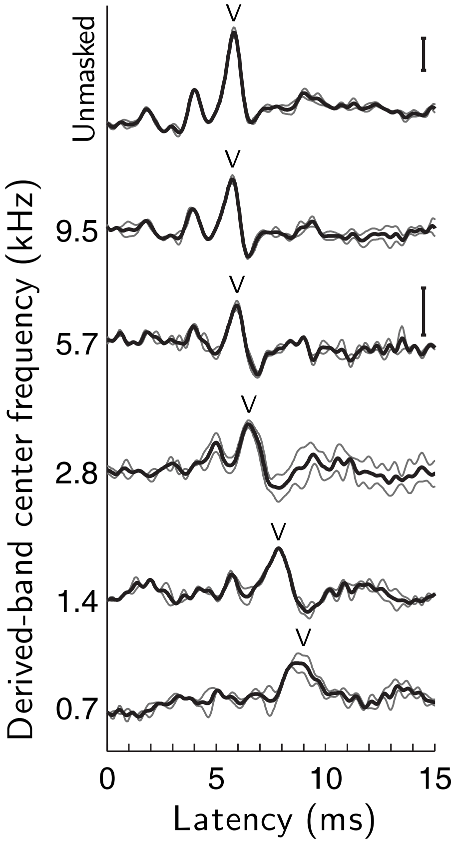

FIG. 1. Examples of unmasked and derived-band ABRs to 93-dB ppe SPL clicks from one listener. Two replications (gray) and their average (black) are shown. Wave *V*s are indicated by the corresponding symbols. The bars to the right represent 200 nV. If no bar is shown, the nearest bar above holds.

Elberling (1994), was used for estimation of the auditory evoked potentials. Two replications, each consisting of 4096 sweeps, were recorded. The 4096 sweeps were subdivided into 16 equally sized blocks and averaged. Each block was weighted in inverse proportion to its amount of background noise, which was estimated as the sweep-to-sweep variance at a single point in time (Elberling and Don, 1984). The residual background noise level in the final evoked potential estimates was 23 nV, averaged across listeners and conditions.

4. Analysis

Narrow-band cochlear contributions to the ABR were derived by means of the derived-band technique (e.g., Don and Eggermont, 1978; Parker and Thornton, 1978b, 1978a). Derived-band ABRs, i.e., differences between the ABRs to clicks presented in adjacent high-pass maskers, were obtained and the corresponding wave-*V* latencies were extracted. The center frequencies of the derived bands were computed as the geometric means of the two corresponding high-pass cut-off frequencies (Parker and Thornton, 1978a). The frequency of 11.3 kHz, where the acoustic-click power was attenuated by 30 dB, was chosen as the upper frequency limit of the highest derived band. Hence, the following frequencies were assigned to the derived bands: 0.7, 1.4, 2.8, 5.7, and 9.5 kHz.

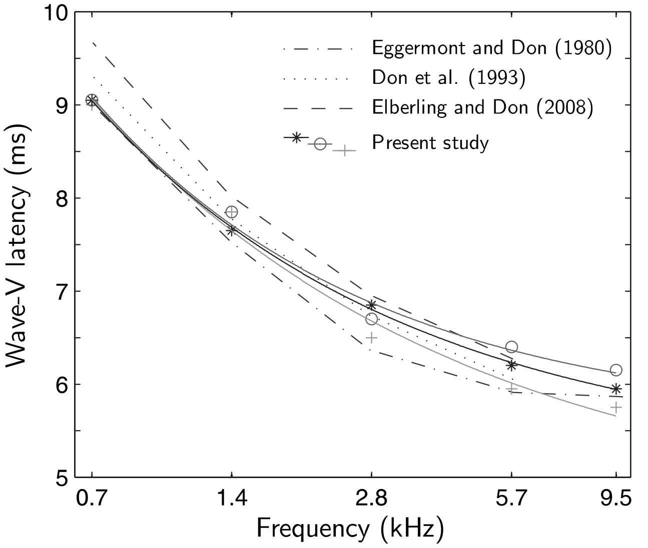

FIG. 2. Measured derived-band ABR wave-*V* latencies (symbols) for three listeners in response to 93-dB ppe SPL clicks, as a function of the derived-band center frequency. The solid curves show individual model fits according to Eq. (1). For comparison, the dash-dotted, dotted, and dashed curves show latency results from Eggermont and Don (1980), Don *et al.* (1993), and Elberling and Don (2008), respectively. The same center frequencies as in Elberling and Don (2008) were assigned to the derived-band latencies of Eggermont and Don (1980), since the same high-pass masking noise stimuli were used in both studies.

Figure 1 illustrates a series of derived-band ABRs from one listener. Wave *V*s are indicated. As can be seen, wave-*V* latencies increased with decreasing derived-band center frequency. For the further analysis of the wave-*V* latencies, the following latency model was adapted from Neely *et al.* (1988):

$$\tau(f) = a + bf^{-d}, \tag{1}$$

where f represents the derived-band center frequency, normalized to 1 kHz, and a, b, and d are fitting constants. The model parameter a represents an asymptotic delay. It reflects the post-cochlear contributions, i.e., synapse and neural conduction delays, to the wave-*V* latency, which are independent of frequency (cf. Don and Eggermont, 1978; Ponton *et al.*, 1992; Ruggero, 1992).

B. Results

Figure 2 shows the measured (symbols) and fitted (solid curves) wave-*V* latencies. The results of all three listeners were similar. Latencies decreased with increasing frequency from about 9 ms at 0.7 kHz to about 6 ms at 9.5 kHz. For comparison, previously reported latencies from Eggermont and Don, 1980 (dash-dotted curve), Don *et al.*, 1993 (dotted curve), and Elberling and Don, 2008 (dashed curve) are shown. The results of the present study agree well with those from the earlier studies. The latency model specified in Eq. (1) provided a good description of the individual latency data, with a residual root-mean-square (rms) fitting error of 0.09 (0.03, 0.13) ms, averaged across listeners (values in parentheses represent the range of individual results). The mean estimated parameters were a=5.1 ms, b=3.2 ms, and d=0.6.

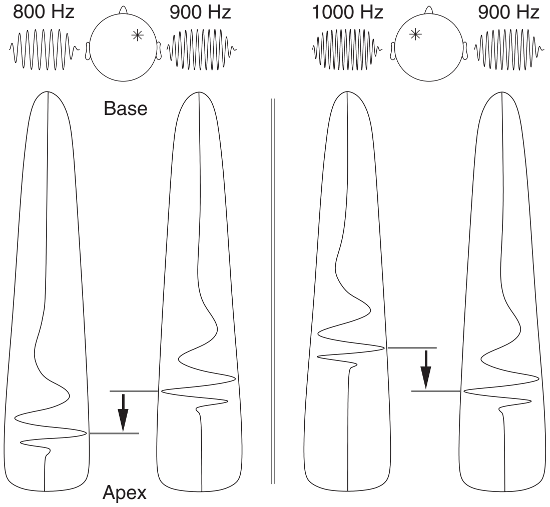

FIG. 3. Sketch of the stimuli used in the lateralization task, for the 800|900-Hz (top left) and 1000|900-Hz (top right) conditions. In the depicted configuration, the left ear corresponds to the ABR test-ear. Basilar membrane traveling waves are indicated at the bottom. It is assumed that the CRT disparities, indicated by the arrows, can be measured in terms of the ITDs that center the percepts at the midline.

III. LATERALIZATION OF MISMATCHED TONES

A. Method

1. Listeners

The lateralization measurements were performed by the same listeners who participated in the ABR measurements.

2. Stimuli and procedures

Short trains of tone bursts with interaurally mismatched frequencies f_1 and f_2 were presented to the two ears, as illustrated in Fig. 3. In the following, the notation $f_1|f_2$ is used where f_1 represents the frequency of the tone presented in the ABR test-ear and f_2 the frequency of the tone presented in the other ear. The considered tone frequencies were 400|480, 800|900, 1000|900, and 1400|1550 Hz. Each tone burst had a total duration of 40 ms, including an exponential onset with a rise time of 10 ms and a 10-ms raised-cosine shaped offset-ramp. In contrast to Scharf *et al.* (1976) and Buus *et al.* (1984), who used exponential ramps at onset and offset, a cosine offset-ramp was used here in order to minimize spectral splatter. The tones were presented in sine phase, i.e., the onset-ramp started with a positive-going zero crossing of the sinusoid. Each train consisted of six tone bursts, separated by 40-ms silent gaps. Its lateralization was varied by introducing a waveform delay to one of the ears, giving rise to an ITD. The ITD that produced a unified percept centered at the midline was measured.

A two-interval, two-alternative, forced-choice task was used. The first interval always contained the diotic reference tone-burst train, consisting of both tones (with frequencies f_1 and f_2) in both ears, while the second interval contained the $f_1|f_2$ target train. Listeners were instructed to indicate if the latter was lateralized to the left or right side relative to the reference train. In order to ease the task, the whole trial consisting of reference and target train was repeated once before the listener made a response. If the target train was lateralized to the right, the ITD was adjusted such that the percept would move further to the left in the next presentation, and vice versa. Following the adaptive procedure for subjective judgments introduced by Jesteadt (1980), two sequences of trials were interleaved, tracking 71% (one up, two down) and 29% (two up, one down) lateralization to the right. Each of these sequences was terminated after ten reversals, and the tracked ITDs were estimated as the arithmetic means of all ITD values following the sixth reversals. Subsequently, the ITD yielding a centered percept was estimated by calculating the mean of the two ITDs leading to 71% and 29% lateralization judgments to the right.

ITDs were measured for tone levels of 50 and 75 dB SPL. In addition to the ITDs in quiet, for the 800|900-Hz tones at 75 dB, ITDs were measured in the presence of a diotic notched-noise background (flat-spectrum noise bands of 100–700 and 1000–9000 Hz), which limited spread of excitation. The noise was presented continuously during the whole run, with a spectrum level of 16 dB SPL. For higher levels, a fused position of the tones could no longer be perceived.

Prior to actual data collection, listeners received up to ten runs of training until consistent ITD results were obtained. The final ITD was estimated as the arithmetic mean over four interleaved runs. If the standard deviation (SD) over these runs, relative to the mean ITD, exceeded a factor of 0.1, additional runs were taken and the average of all was used. The final relative standard error of the ITD estimate, averaged across listeners and conditions, was 0.05.

3. Loudness balancing

In addition to the conditions where the tones were presented at equal SPLs, ITDs were measured with the tones balanced in loudness between the two ears. Loudness balancing was also applied by Zerlin (1969). The adaptive procedure introduced by Jesteadt (1980) was used for the loudness balancing of the frequency-mismatched tones. The first interval contained the f_1-tone, presented to the ABR test-ear, and the second interval contained the f_2-tone, presented to the other ear. Listeners were instructed to indicate if the second tone was perceived as softer or louder than the first tone. As in the lateralization task, the whole trial was repeated once before the listener made a response. The interaural level balance was adjusted to yield both 71% and 29% judgments of the second tone to be the louder one. The point of equal loudness was estimated as the arithmetic mean of these two loudness adjustments (in decibels). An equal number of runs were performed with the opposite order of presentation, i.e., with the f_2-tone presented in the first interval and the f_1-tone presented in the second interval.

The final level adjustment for loudness balancing was estimated as the arithmetic mean over at least six interleaved runs. The final standard error of the level adjustment was 0.4 (SD 0.2) dB, averaged across listeners and conditions. There were no significant differences between listeners and conditions $[p > 0.1]$.

TABLE I. The ITDs yielding centered percepts of the tones with interaurally mismatched frequencies f_1 and f_2, for three listeners (the numbers in parentheses represent standard errors). LB denotes loudness balancing. The ABR wave-V latency differences $\Delta\tau_{\mathrm{ABR}}$ between the frequencies f_1 and f_2 are also given for the individual listeners. The values in square brackets are based on extrapolations beyond the range of measured frequencies. Conditions for which the listener could not perform the lateralization task are indicated by "NM" (not measurable). Dots indicate combinations that were not measured.

Tone level	$f_1 \mid f_2$ (kHz)	ITD (μs) for NH_1		ITD (μs) for NH_2		ITD (μs) for NH_3		$\Delta\tau_{\mathrm{ABR}}$ (μs)		
		With LB	Without LB	With LB	Without LB	With LB	Without LB	NH_1	NH_2	NH_3
50 dB	$0.4 \mid 0.48$	442(38)	340(40)	404(6)	396(8)	357(8)	335(5)	[580	648	597]
	$0.8 \mid 0.9$	205(12)	264(13)	186(6)	232(10)	184(7)	207(3)	261	260	255
	$1.0 \mid 0.9$	187(18)	185(13)	224(2)	262(3)	173(3)	180(9)	220	215	212
	$1.4 \mid 1.55$	110(5)	138(9)	NM	129(22)	NM	NM	167	151	155
	$\substack{0.8\\1.0}\mid$	392(21)	449(18)	410(7)	494(11)	356(8)	387(9)	481	475	467
75 dB	$0.8 \mid 0.9$	99(3)	⋯	79(5)	⋯	61(7)	⋯			
	$0.8 \mid 0.9$ in noise	190(6)	⋯	183(2)	⋯	192(6)	⋯			

4. Apparatus

The stimuli were generated in MATLAB® and converted to analog signals using a 24-bit digital-to-analog converter (RME DIGI96/8) with a sampling rate of 96 kHz. The stimuli were presented in a double-walled sound-attenuating booth via Sennheiser HD580 headphones. Calibrations were done using an ear simulator (IEC 60318-1and -2, 1998; B&K 4153 with flat plate) and, prior to playing, 128-tap linear-phase FIR equalization filters were applied to the stimuli, rendering the headphone frequency response flat.

B. Results and discussion

1. Response-time differences

The results of the lateralization measurements for the three listeners are presented in Table I. It shows the ITDs that led to centered percepts of the 50- and 75-dB tones with interaurally mismatched frequencies f_1 and f_2. The ITDs are given for the conditions with and without interaural loudness balancing. As illustrated in Fig. 3, the frequency-mismatched tones with zero ITD were always lateralized toward the ear receiving the higher-frequency tone, consistent with previous reports in literature (e.g., von Békésy, 1963b; Zerlin, 1969). Hence, the sound presented to this ear required a delay in order to center the percept (for this reason, ITDs are stated only in absolute terms in the following). The centering ITDs were generally consistent and well reproducible. Therefore, the standard errors of the ITD estimates were relatively small. For comparison, the objective ABR wave-V latency differences $\Delta\tau_{\mathrm{ABR}}$ are also represented in Table I (rightmost column). They were calculated on the basis of the individual latency fits to the derived-band ABR data, which followed the model in Eq. (1) and were shown in Fig. 2. The lowest derived-band frequency was 700 Hz. Therefore, the extrapolation to lower frequencies (400|480 Hz) should be regarded with caution. At the remaining frequencies of 800|900, 1000|900, and 1400|1550 Hz (second, third, and fourth rows in Table I, respectively), the perceptual ITD-based measure and the objective ABR-based measure yielded very similar results. The average rms deviation between the ITDs (without loudness balancing) and the latency differences $\Delta\tau_{\mathrm{ABR}}$ was 39 μs. The correspondence between the behavioral and the objective data is remarkably good, given the different experimental paradigms. It strongly supports the hypothesis that the ITDs that produced centered sound images reflected differences in CRTs between remote places on the BM.

The ITDs reflect interaural time differences whereas the ABR latency differences $\Delta\tau_{\mathrm{ABR}}$ reflect monaural time differences. Hence, part of the remaining deviations between these two could be due to differences in CRTs between the left and right cochleae (e.g., differences in the cochlear frequency-place maps). Therefore, the ITDs for the 800|900-Hz and 1000|900-Hz tone pairs were added (see fifth row in Table I). Since these tone pairs shared the common reference frequency of 900 Hz (cf. Fig. 3), the sum estimates the time difference between 800 and 1000 Hz in the ABR test-ear alone. Still, similar deviations from the ABR latencies as for the single-tone-pair ITDs were observed for these "monaural" time differences. Hence, the remaining deviations did not seem to be attributable to asymmetries between the left and right cochleae.

In addition to the measurements at 50 dB, for the 800|900-Hz tones, measurements were also performed at the higher tone level of 75 dB. For all listeners, ITDs were shorter at 75 dB than at 50 dB, by an average factor of 2.5. However, in the presence of the notched-noise masker, the ITDs obtained with the 75-dB tones were essentially identical to those obtained with 50-dB tones presented in quiet. This is consistent with the following interpretation in terms of excitation spread on the BM. The higher the tone level, the larger is the spread of excitation toward places with higher CFs than the nominal tone frequencies f_1 and f_2. The disparities in CRT between these places are smaller than at the nominal places due to the exponentially decreasing latency-frequency dependence (cf. Fig. 2). Therefore, smaller centering ITDs would be expected for the higher tone level of 75 dB than for the lower level of 50 dB. The notched noise limits excitation spread. This may explain why similar ITDs were obtained for the 75-dB tones in noise as for the 50-dB tones in quiet, for which spread of excitation plays a minor role. Hence, the observed effects of tone level and background noise further indicate that the perceived lateralization of the mismatched tones reflected cochlear disparities.

Different stimuli, clicks versus tones, were used for the ABR recordings and the lateralization measurements, respectively. It seems reasonable to assume that stimulation at equal sensation levels results in similar levels of neural excitation, summed across the BM. The sensation level of the 93-dB ppe SPL clicks was 59 dB, averaged across listeners. The average sensation level of the mismatched 50-dB SPL tones with center frequencies of 890 and 1470 Hz was 49 dB (the same 3I-3AFC task was used for estimation of the click and tone thresholds). However, the tones excited only a limited part of the cochlea, while the broadband clicks excited most of the cochlea partition. Hence, the "effective" click levels in the one-octave-wide derived bands were lower than the nominal click level. In order to estimate these levels, the portion of the click power falling within the derived bands was calculated based on the acoustic-click power spectra. For the 0.7- and 1.4-kHz derived bands, this yielded values of -11 and -9 dB relative to the broadband click level, respectively. Hence, within these derived bands, the effective click level was about 83 dB ppe SPL, corresponding to a sensation level of 49 dB, which matches the sensation level of the mismatched tones. Also, remaining level differences should be of minor importance, since the 75-dB tones yielded very similar ITDs to the 50-dB tones when notched-noise masking was applied.

All three listeners had more difficulties with the lateralization task for the mid-frequency tones ($1400|1550$ Hz) than for the low-frequency tones. At $1400|1550$ Hz, listener NH_2 could not consistently lateralize the mismatched tones when loudness balancing was applied, while listener NH_3 could not consistently lateralize the tones whether loudness balancing was applied or not. None of the listeners could perform the task reliably for frequencies above 1.5 kHz. Here, the sound image could not be lateralized with reasonable precision. It was perceived as rather diffuse and often did not cross the midline.

2. Loudness balancing

For all tone pairs, ITDs changed systematically when loudness balancing was applied: The ITD increased (decreased) when the level of the higher-frequency tone was increased (decreased). The level adjustment was 0.7 (SD 0.4) dB, averaged across listeners and conditions, without showing a systematic pattern across listeners and conditions. The ITDs obtained without loudness balancing seemed to match the objective latency differences $\Delta\tau_{\mathrm{ABR}}$ slightly better than the ones obtained with loudness balancing. The average rms deviations were 39 and 66 μs, respectively, excluding the $400|480$-Hz data.

Depending on the mechanism underlying the lateralization of the mismatched tones, loudness or level imbalances could influence the results of the lateralization measurements. While a temporal (phase-locking-based) mechanism should hardly be affected, a mechanism based on interaural level differences should be sensitive to level/loudness imbalances. The observed systematic change in ITDs with loudness balancing may indicate that interaural level cues contributed to the lateralization of the mismatched tones, although the small ITD changes might simply reflect changes in CRT with tone level. In any case, the centering ITDs obtained with and without loudness balancing were fairly comparable. This is consistent with the hypothesis that the lateralization was, at least to some extent, based on a temporal mechanism. This hypothesis is corroborated by the finding that tone-onset phase influences the lateralization of mismatched tones for frequencies below about 2 kHz (Scharf *et al.*, 1976; Buus *et al.*, 1984). With respect to the estimation of CRT disparities, the observed invariance of the results to loudness balancing is crucial. If the lateralization depended strongly on interaural level (or loudness) imbalances, it would be impossible to assess disparities in CRTs with this method. The trade-off between timing and level would give rise to unresolvable ambiguities.

As mentioned above, the behavioral results obtained without loudness balancing matched the objective data better than the ones obtained with loudness balancing. The observed loudness imbalances might have been due to different amounts of excitation or specific loudness (Moore *et al.*, 1997) at the two tone frequencies f_1 and f_2 as well as due to frequency-independent *between-ear* differences in excitation or specific loudness. Depending on which of these two factors was dominant, either equal loudness or equal SPLs at the two ears would be more appropriate for the lateralization paradigm. Frequency-independent between-ear differences in excitation/specific loudness would affect the lateralization of the diotic reference stimulus and mismatched target stimulus in the same way. Therefore, the ITD necessary for matching their positions would not be affected, as long as loudness balancing was applied neither to the reference stimulus nor to the target stimulus.[1] Hence, the better match between the objective results and the behavioral results obtained without loudness balancing and the lack of a systematic effect of tone frequency on loudness balancing suggest that the observed loudness imbalances may have reflected between-ear differences rather than frequency-dependent variations in excitation/specific loudness.

3. Traveling–wave velocity

Assuming that the centering ITDs and latency differences $\Delta\tau_{\mathrm{ABR}}$ reflected travel times on the BM, the corresponding TW velocities were estimated using the cochlear frequency-place map supplied by Greenwood (1961).[2] The ratio of the distance between the CF places (corresponding to the mismatched-tone frequencies) and the centering ITD was taken as behavioral estimate of TW velocity at the geometric mean frequency of the two tones. Objective ABR-based estimates were derived by substituting the frequency f in Eq. (1) by the CF place x on the BM [Eq. (2) in Greenwood, 1961] and taking the derivative $-dx/d\tau$ as estimate of velocity.

Figure 4 shows the TW velocity estimates, based on the ABR latencies (curves) and the centering ITDs (bullets) obtained for the 50-dB tones without loudness balancing. The behavioral velocity estimate at 890 Hz (geometric means of 800 and 1000 Hz) is based on the "monaural" time difference

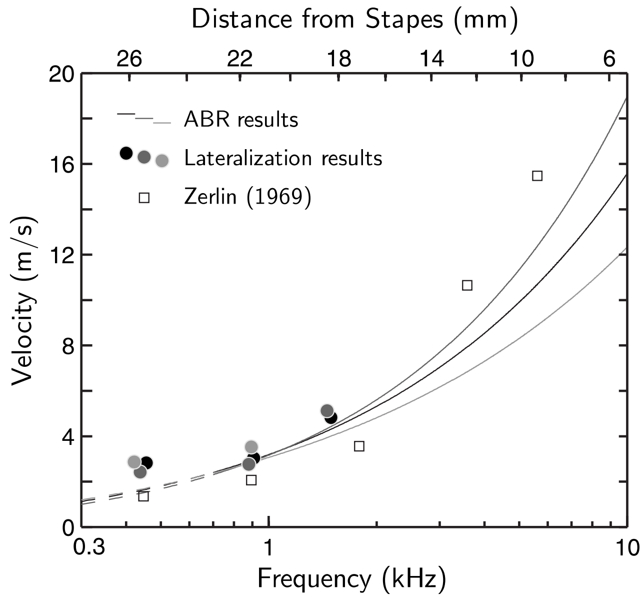

FIG. 4. TW velocity as a function of frequency/distance from stapes for three listeners. The solid curves represent the individual velocity estimates derived from the derived-band ABR latencies. At low frequencies, the curves are dashed since they are extrapolated beyond the actual measurement range. The bullets denote the estimates based on the mismatched-tone ITDs. For better visibility, they are slightly horizontally displaced for the individual listeners. The squares are corresponding estimates based on the ITDs reported by Zerlin (1969).

obtained by summing the 800|900 and 1000|900-Hz ITDs. For direct comparison, the open squares indicate velocities that were derived from Zerlin's (1969) ITDs.[3]

The ITD-based velocity estimates were consistent with the ABR-based velocity estimates. In both measures, velocities increased with increasing frequency. In order to compare the ITD-based estimates at 440 Hz with the ABR-based estimates, the ABR data were extrapolated beyond the actual measurement range (dashed part of the curves). Here, the deviations between the two measures were larger than at the higher frequencies of 890 and 1470 Hz, reflecting the corresponding deviations of the CRT estimates (compare ITDs and $\Delta\tau_{\mathrm{ABR}}$ values in Table I). The larger behavioral velocity estimates at 440 Hz might indicate that the actual latency-frequency functions were less steep at the low frequencies (below about 700 Hz) than the predictions based on the extrapolation of the ABR latencies (Fig. 2). This would be consistent with the latency-frequency curves in Fig. 1 of Neely *et al.* (1988), obtained from tone-burst-evoked ABRs, which showed shallower slopes for frequencies below about 500 Hz than for the higher frequencies. However, Ruggero and Temchin (2007) noted that the use of different tone-burst rise times for the different frequencies in the study by Neely *et al.* (1988) could have affected the observed ABR latencies.

Only small inter-individual differences were observed for frequencies up to 2 kHz, consistent with Donaldson and Ruth (1993). For frequencies above 1.5 kHz, no centering ITDs and thus no behavioral velocity estimates could be obtained in this study. At low frequencies, the velocity estimates were higher than the ones based on Zerlin's (1969) ITDs (open squares).[4] Zerlin (1969) also estimated TW velocities at high frequencies. These velocities were larger than the velocities at low frequencies and roughly consistent with the present ABR-based estimates.

IV. LIMITATIONS OF THE LATERALIZATION PARADIGM

A. Critical band and lateralization threshold

Despite the encouraging results of the lateralization paradigm for tone frequencies up to 1.5 kHz, no behavioral estimates of CRT could be obtained at higher frequencies. This was due to fundamental limitations in the lateralization paradigm, which are discussed in the following. In principle, a large frequency mismatch $|f_2 - f_1|$ between the tones would be desirable to increase the accuracy of the ITD estimate. However, with increasing frequency mismatch, it becomes increasingly difficult to attribute a fused position (Scharf, 1972). More importantly, the lateralization threshold, i.e., the ITD for which the position of a non-centered sound object can just be distinguished from that of a centered object, increases strongly as soon as the interaural frequency mismatch exceeds a value that corresponds to the critical bandwidth for that frequency (Scharf *et al.*, 1976; Buus *et al.*, 1984). Scharf *et al.* (1976) found this bandwidth to be roughly independent of tone level and tone duration. The centering ITD, reflecting CRT disparity, needs to be larger than the corresponding lateralization threshold in order to be measurable. Therefore, in the present study, each tone pair was chosen such that the frequency mismatch between the tones did not exceed the critical bandwidth at the corresponding center frequency. The tone level of 50 dB SPL should have been comparable to the levels used by Zerlin (1969), which corresponded to an approximate loudness level of 50 phon. It was chosen as a compromise between decreasing lateralization thresholds and increasing spread of excitation with increasing tone level.

The feasibility of the measurements can, in principle, be predicted by comparing expected CRT disparities for maximally mismatched tones (tones that fall just within the same critical band) with the corresponding lateralization thresholds. As mentioned above, the CRT disparity for the mismatched tones can only be measured in terms of the ITD that is required to center the percept, if this ITD is larger than the lateralization threshold (which determines ITD sensitivity). This is discussed in the following.

B. Predicted CRT disparities from objective data

Critical bandwidths at 500, 1000, 2000, 4000, and 6000 Hz were extracted by digitizing the figures in Scharf *et al.* (1976). The obtained values were 115, 163, 310, 702, and 1080 Hz, respectively. In the next step, the frequencies of maximally mismatched tones were calculated such that the geometric means of the two frequencies were equal to 500, 1000, 2000, 4000, and 6000 Hz. At 2000 Hz, for example, the tone frequencies were 1850 and 2160 Hz. Distances between the corresponding CF places on the BM were calculated according to the Greenwood (1961) frequency-place map. Next, objective TW velocity estimates from different studies, as given in Fig. 10 of Donaldson and Ruth (1993),

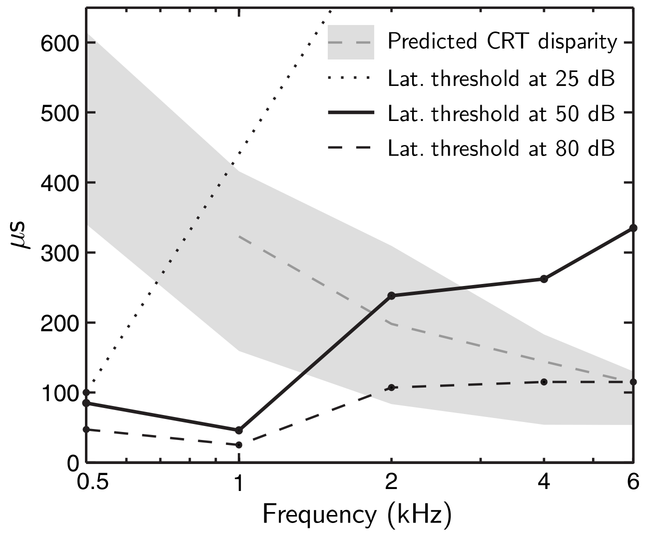

FIG. 5. Predicted CRT disparities (gray) for maximally mismatched tones as a function of the center frequency of the tones. The gray shaded area indicates CRT disparities based on the TW velocity estimates given in Donaldson and Ruth (1993). The gray dashed curve shows disparity estimates based on the TW velocity estimates obtained in the present study (curves in Fig. 4). The black curves indicate the lateralization thresholds at 25 (dotted curve), 50 (solid curve), and 80 dB SPL (black dashed curve), obtained by Scharf *et al.* (1976) and Buus *et al.* (1984).

were used to predict CRT disparities ("travel times") corresponding to these distances along the BM. The different velocity estimates yielded a range of CRT disparities, which are shown as gray shaded area in Fig. 5. The gray dashed curve indicates disparity estimates that are based on the average TW velocities obtained for the three listeners of the present study (curves in Fig. 4). As can be seen, CRT disparities for the maximally mismatched tones decrease with increasing center frequency of the tones. Furthermore, the estimates based on TW velocities obtained in the present study are consistent with those based on the TW velocities in Donaldson and Ruth (1993). Figure 5 also shows the lateralization thresholds at the different tone levels of 25 (dotted curve), 50 (solid curve), and 80 dB SPL (black dashed curve), obtained by Scharf *et al.* (1976) and Buus *et al.* (1984).[5] Up to frequencies of about 1.5 kHz, the predicted CRT disparities are larger than the corresponding lateralization thresholds for mismatched 50-dB tones (solid curve) and are therefore measurable. However, with increasing frequency, the CRT disparities fall below the lateralization thresholds and are not measurable at a tone level of 50 dB. In theory, they are measurable using tone levels of about 80 dB and higher, since lateralization thresholds are smaller at these higher levels (black dashed curve). However, this assumes that spread of excitation can be adequately limited, for example, by means of notched-noise masking. For high frequencies of about 4 kHz or higher, the predicted CRT disparities are too small to be measurable, even for tone levels of 80 dB. These predictions are consistent with the finding from the present study that ITDs could not be obtained for 50-dB tones at frequencies above 1.5 kHz.

C. Comparison with Zerlin's study

The frequency mismatches for all tone pairs used by Zerlin (1969) exceeded the critical bandwidths given by Scharf *et al.* (1976) and Buus *et al.* (1984). For the 3200|4000-Hz and 5000|6300-Hz tone pairs, the reported centering ITDs clearly fall below the corresponding lateralization thresholds in those studies. Furthermore, Scharf *et al.* (1976) emphasized the importance of controlled tone-onset phases for tone frequencies below about 2 kHz: Without controlling the onset phase, their ITD data were inconsistent and the observed lateralization thresholds became substantially larger. Zerlin, however, did not control onset phases. Hence, the validity of his results appears questionable both at low and high frequencies.

One might argue that part of the discrepancies could be due to different ramp durations. Zerlin (1969) used 2.5-ms ramps, whereas 10-ms ramps were used in the present study as well as in Scharf *et al.* (1976) and Buus *et al.* (1984). However, even with such short ramp durations (tested in pilot measurements), it was not possible to obtain consistent ITD data at high frequencies. Apart from this, the percept gained a click-like character indicating a loss of frequency specificity.

V. CONCLUSIONS

For frequencies up to 1.5 kHz, the lateralization of mismatched tones yielded estimates of CRT disparities (across remote BM places) and TW velocities that were reasonably accurate and consistent with objective estimates based on ABR measurements. However, due to intrinsic limitations of the lateralization paradigm, it was impossible to obtain behavioral estimates of CRT disparities at high frequencies.

Besides the possibility of studying aspects of the spatiotemporal BM response pattern other than CRT (e.g., response amplitude), a further step could be to investigate relations between individual estimates of CRT (disparities) and performance in other psychoacoustic tasks that have been discussed in the context of spatiotemporal processing (e.g., pitch perception and tone-in-noise detection). Here, the inclusion of hearing-impaired listeners may be crucial. Alterations in the spatiotemporal BM response, due to hearing impairment, might result in reduced performance in these tasks compared to normal-hearing listeners. The larger-than-normal across-listener variability within the hearing-impaired population may allow the study of such relations. Furthermore, it would be interesting to model the effect of CRT alterations in the framework of spatiotemporal models (cf. Carney, 1994). The lateralization method presented in this study might provide valuable information about such CRT alterations, particularly at low frequencies (below 500 Hz), where the accuracy of objective methods is limited.

ACKNOWLEDGMENTS

The authors wish to thank Dimitrios Christoforidis and Dr. James Harte for technical support and valuable scientific discussions about various aspects of this project. They also thank Brent C. Kirkwood, Torben Poulsen, Brenda L. Lonsbury-Martin, and two anonymous reviewers for their valuable comments on an earlier version of this manuscript. They are grateful to the listeners for their participation in testing. Part of this work was supported by the Danish Research Foundation, the Danish Graduate school SNAK "Sense organs, neural networks, behavior, and communication," and the Oticon Foundation.

[1] The reference stimulus was not balanced in loudness since, for matched-frequency tones, equal SPLs instead of equal loudness at the two ears would give rise to a percept centered at the midline. As discussed by Durlach *et al.* (1981), the binaural system adapts to between-ear gain differences in such a way that equal-SPL tones are perceived at the midline. In this way, the correlation of auditory perception with visual and tactile perceptions is maximized.

[2] The further assumption is made that the response time at a given CF place of the BM is the same for tonal stimulation with frequencies at and below this CF. This corresponds to constant group delays, i.e., constant slopes of the BM phase response (cf. Ruggero and Rich, 1987; Robles and Ruggero, 2001). For the mismatched 800|900-Hz tone pair, for example, the TWs in response to the 800- and 900-Hz tones would reach the 900-Hz CF place at the same time. Hence, CRT differences would reflect the travel time between the 800- and 900-Hz CF places.

[3] Zerlin (1969) used the cochlear frequency-place map supplied by von Békésy (1963a) in order to derive TW velocities from the centering ITDs. In the present study, the Greenwood (1961) cochlear map was taken as a basis of all TW velocity estimates. Therefore, the TW velocities shown here were derived directly from the ITDs reported by Zerlin (1969) using the Greenwood (1961) map.

[4] These deviations cannot be attributed to the fact that Zerlin (1969) applied loudness balancing. Velocity estimates based on the ITDs obtained with loudness balancing (not shown) always fell in the same range or above the ones obtained without loudness balancing, but never below as Zerlin's.

[5] In the data by Buus *et al.* (1984) the actual tone level at 500 Hz was 59 dB SPL, not 50 dB SPL.

Buss, E., Hall, J. W., and Grose, J. H. (**2004**). "Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss," Ear Hear. **25**, 242–250.

Buus, S., Scharf, B., and Florentine, M. (**1984**). "Lateralization and frequency selectivity in normal and impaired hearing," J. Acoust. Soc. Am. **76**, 77–86.

Carney, L. H. (**1994**). "Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness," Hear. Res. **76**, 31–44.

Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., and Colburn, H. S. (**2002**). "Auditory phase opponency: A temporal model for masked detection at low frequencies," Acta. Acust. Acust. **88**, 334–347.

Clarey, J. C., Barone, P., and Imig, T. J. (**1992**). "Physiology of thalamus and cortex," in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer-Verlag, New York), pp. 232–334.

Deatherage, B. H., and Hirsh, I. J. (**1959**). "Auditory localization of clicks," J. Acoust. Soc. Am. **31**, 486–492.

Deng, L., and Geisler, C. D. (**1987**). "A composite auditory model for processing speech sounds," J. Acoust. Soc. Am. **82**, 2001–2012.

Don, M., and Eggermont, J. J. (**1978**). "Analysis of the click-evoked brainstem potentials in man using high-pass noise masking," J. Acoust. Soc. Am. **63**, 1084–1092.

Don, M., and Elberling, C. (**1994**). "Evaluating residual background noise in human auditory brain-stem responses," J. Acoust. Soc. Am. **96**, 2746–2757.

Don, M., Ponton, C. W., Eggermont, J. J., and Masuda, A. (**1993**). "Gender differences in cochlear response time: An explanation for gender amplitude differences in the unmasked auditory brain-stem response," J. Acoust. Soc. Am. **94**, 2135–2148.

Donaldson, G. S., and Ruth, R. A. (**1993**). "Derived band auditory brainstem response estimates of traveling wave velocity in humans. I: Normal-hearing subjects," J. Acoust. Soc. Am. **93**, 940–951.

Durlach, N. I., Thompson, C. L., and Colburn, H. S. (**1981**). "Binaural interaction in impaired listeners. A review of past research," Audiology **20**, 181–211.

Eggermont, J. J. (**1976**). "Analysis of compound action potential responses to tone bursts in the human and guinea pig cochlea," J. Acoust. Soc. Am. **60**, 1132–1139.

Eggermont, J. J., and Don, M. (**1980**). "Analysis of the click-evoked brainstem potentials in humans using high-pass noise masking. II. Effect of click intensity," J. Acoust. Soc. Am. **68**, 1671–1675.

Elberling, C., and Don, M. (**1984**). "Quality estimation of averaged auditory brainstem responses," Scand. Audiol. **13**, 187–197.

Elberling, C., and Don, M. (**2008**). "Auditory brainstem responses to a chirp stimulus designed from derived-band latencies in normal-hearing subjects," J. Acoust. Soc. Am. **124**, 3022–3037.

Elberling, C., and Wahlgreen, O. (**1985**). "Estimation of auditory brainstem response, ABR, by means of Bayesian inference," Scand. Audiol. **14**, 89–96.

Gorga, M. P., Kaminski, J. R., Beauchaine, K. A., and Jesteadt, W. (**1988**). "Auditory brainstem responses to tone bursts in normally hearing subjects," J. Speech Hear. Res. **31**, 87–97.

Greenwood, D. D. (**1961**). "Critical bandwidth and the frequency coordinates of the basilar membrane," J. Acoust. Soc. Am. **33**, 1344–1356.

IEC 60318-1 (**1998**). "Electroacoustics—Simulators of human head and ear—Part 1: Ear simulator for the calibration of supra-aural earphones," International Electrotechnical Commission, Geneva.

IEC 60318-2 (**1998**). "Electroacoustics—Simulators of human head and ear—Part 2: An interim acoustic coupler for the calibration of audiometric earphones in the extended high-frequency range," International Electrotechnical Commission, Geneva.

IEC 60711 (**1981**). "Occluded ear simulator for the measurement of earphones coupled to the ear by ear inserts," International Electrotechnical Commission, Geneva.

ISO 389-8 (**2004**). "Acoustics—Reference zero for the calibration of audiometric equipment—Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones," International Organization for Standardization, Geneva.

Jesteadt, W. (**1980**). "An adaptive procedure for subjective judgments," Percept. Psychophys. **28**, 85–88.

Joris, P. X., de Sande, B. V., Louage, D. H., and van der Heijden, M. (**2006**). "Binaural and cochlear disparities," Proc. Natl. Acad. Sci. U.S.A. **103**, 12917–12922.

Loeb, G. E., White, M. W., and Merzenich, M. M. (**1983**). "Spatial cross-correlation. A proposed mechanism for acoustic pitch perception," Biol. Cybern. **47**, 149–163.

Magezi, D. A., and Krumbholz, K. (**2008**). "Can the binaural system extract fine-structure interaural time differences from noncorresponding frequency channels?," J. Acoust. Soc. Am. **124**, 3095–3107.

Moore, B. C. J. (**1996**). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," Ear Hear. **17**, 133–161.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). "A model for the prediction of thresholds, loudness, and partial loudness," J. Audio Eng. Soc. **45**, 224–240.

Moore, B. C. J., and Skrodzka, E. (**2002**). "Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation," J. Acoust. Soc. Am. **111**, 327–335.

Neely, S. T., Norton, S. J., Gorga, M. P., and Jesteadt, W. (**1988**). "Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli," J. Acoust. Soc. Am. **83**, 652–656.

Norton, S. J., and Neely, S. T. (**1987**). "Tone-burst-evoked otoacoustic emissions from normal-hearing subjects," J. Acoust. Soc. Am. **81**, 1860–1872.

Parker, D. J., and Thornton, A. R. (**1978a**). "Frequency specific components of the cochlear nerve and brainstem evoked responses of the human auditory system," Scand. Audiol. **7**, 53–60.

Parker, D. J., and Thornton, A. R. (**1978b**). "The validity of the derived cochlear nerve and brainstem evoked responses of the human auditory system," Scand. Audiol. **7**, 45–52.

Ponton, C. W., Eggermont, J. J., Coupland, S. G., and Winkelaar, R. (**1992**). "Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: Evidence from derived auditory brain-stem responses (ABRs)," J. Acoust. Soc. Am. **91**, 1576–1586.

Richter, U., and Fedtke, T. (**2005**). "Reference zero for the calibration of audiometric equipment using 'clicks' as test signals," Int. J. Audiol. **44**, 478–487.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Ruggero, M. A. (**1992**). "Physiology and coding of sound in the auditory nerve," in *The Mammalian Auditory Pathway: Neurophysiology*, edited by A. N. Popper and R. R. Fay (Springer-Verlag, New York), pp. 34–93.

Ruggero, M. A. (**1994**). "Cochlear delays and traveling waves: Comments on 'Experimental look at cochlear mechanics'," Audiology **33**, 131–142.

Ruggero, M. A., and Rich, N. C. (**1987**). "Timing of spike initiation in cochlear afferents: Dependence on site of innervation," J. Neurophysiol. **58**, 379–403.

Ruggero, M. A., and Temchin, A. N. (**2007**). "Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods," J. Assoc. Res. Otolaryngol. **8**, 153–166.

Scharf, B. (**1972**). "Frequency selectivity and sound localization," in *Symposium on Hearing Theory*, edited by B. L. Cardozo (IPO, Eindhoven, Germany), pp. 115–122.

Scharf, B., Florentine, M., and Meiselman, C. (**1976**). "Critical band in auditory lateralization," Sens. Processes **1**, 109–126.

Schubert, E. D., and Elpern, B. S. (**1959**). "Psychophysical estimate of the velocity of the traveling wave," J. Acoust. Soc. Am. **31**, 990–994.

Shamma, S., and Klein, D. (**2000**). "The case of the missing pitch templates: How harmonic templates emerge in the early auditory system," J. Acoust. Soc. Am. **107**, 2631–2644.

Shamma, S. A., Shen, N. M., and Gopalaswamy, P. (**1989**). "Stereausis: Binaural processing without neural delays," J. Acoust. Soc. Am. **86**, 989–1006.

Tognola, G., Grandori, F., and Ravazzani, P. (**1997**). "Time-frequency distributions of click-evoked otoacoustic emissions," Hear. Res. **106**, 112–122.

von Békésy, G. (**1933**). "Über den Knall und die Theorie des Hörens (Clicks and the theory of hearing)," Phys. Z. **34**, 577–582.

von Békésy, G. (**1963a**). "Hearing theories and complex sounds," J. Acoust. Soc. Am. **35**, 588–601.

von Békésy, G. (**1963b**). "Three experiments concerned with pitch perception," J. Acoust. Soc. Am. **35**, 602–606.

Yost, W. A. (**1974**). "Discriminations of interaural phase differences," J. Acoust. Soc. Am. **55**, 1299–1303.

Zerlin, S. (**1969**). "Traveling-wave velocity in the human cochlea," J. Acoust. Soc. Am. **46**, 1011–1015.