ELSEVIER

Hearing Research 141 (2000) 57–79

HEARING RESEARCH

www.elsevier.com/locate/heares

Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements☆

M.S. John, T.W. Picton *

Rotman Research Institute, Baycrest Centre for Geriatric Care, University of Toronto, 3560 Bathurst Street, Toronto, Ont. M6A 2E1, Canada

Received 17 August 1999; received in revised form 11 November 1999; accepted 11 November 1999

Abstract

Human auditory steady-state responses were recorded to four stimuli, with carrier frequencies (f_c) of 750, 1500, 3000 and 6000 Hz, presented simultaneously at 60 dB SPL. Each carrier frequency was modulated by a specific modulation frequency (f_m) of 80.6, 85.5, 90.3 or 95.2 Hz. By using four different recording conditions we obtained responses for all permutations of f_m and f_c. The phase delays (P) of the responses were unwrapped and converted to latency (L) using the equation: $L = P/(360 \times f_m)$. The number of cycles of the stimulus that occurred prior to the recorded response was estimated by analyzing the effect of modulation frequency on the responses. These calculations provided latencies of 20.7, 17.7, 16.1 and 16.1 ms for carrier frequencies 750, 1500, 3000 and 6000 Hz. This latency difference of about 4.5 ms between low and high carrier frequencies remained constant over many different manipulations of the stimuli: faster modulation rates (150–190 Hz), binaural rather than monaural presentation, different intensities, stimuli presented alone or in conjunction with other stimuli, and modulation frequencies that were separated by as little as 0.24 Hz. This frequency-related delay is greater than that measured using transient evoked potentials, most likely because of differences in how transient and steady-state responses are generated and how their latencies are determined.

Key words: Auditory evoked potential; Steady-state response; Apparent latency; Phase; Travelling wave

1. Introduction

Physiological responses to auditory stimuli of high frequency generally occur at earlier latencies than responses to stimuli of low frequency. One reason for this is that auditory stimuli initiate a 'travelling wave' in the basilar membrane of the cochlea (Von Békésy, 1960). This travelling wave distributes high frequencies of sound to the regions of the basilar membrane near the stapes and low frequencies to the regions near the apex of the cochlea. As well as thus mediating a frequency-to-place coding in the auditory system, the travelling wave also involves a delay since time is taken to move along the basilar membrane. This frequency-related delay or 'transport time' is the subject of some controversy (Dancer, 1992). Most evidence suggests that the delay is in the millisecond range, increases exponentially with increasing distance along the basilar membrane, and depends mainly upon the passive properties of the basilar membrane, being equivalent for living and dead preparations (Ruggero, 1994).

Another reason for frequency-related delays in the neurophysiological responses is the filtering that occurs during the transduction of the acoustic energy in the stimulus to electrical impulses in the auditory nerve fibers (Eggermont, 1979a). The more sharply a filter is tuned, the longer it takes for the output of the filter to reach maximum amplitude. Since the neural response is delayed by some fraction of this rise time, the neural latency has a 'filter delay' added to the delay due to the transport time. The filter-related delay varies inversely with the frequency to which the filter is tuned. In the normal cochlea both the transport time and the filter

* Corresponding author. Tel.: +1 (416) 785-2500 ext. 3509; Fax: +1 (416) 785-2862; E-mail: picton@psych.toronto.edu

☆ STEADY-STATE software available at http://www.rotman-baycrest.on.ca/users/sasha_j/Masterdl.htm

0378-5955/00/$ – see front matter
PII: S0378-5955(99)00209-9

delay (often considered together as the 'travelling wave delay') contribute to the frequency-related delays of the physiological responses.

The latencies obtained from physiological measurements are sometimes converted to 'travelling wave velocities' (e.g., Parker and Thornton, 1978; Kim et al., 1994) on the basis of the normal length of the basilar membrane and the hypothesized normal locations of frequency-specific regions on this membrane (Von Békésy, 1963). Because the length of the human basilar membrane varies from subject to subject, for example being longer for males than females (Sato et al., 1991), and since the exact locations of the different frequency regions on this membrane are not accurately known, it seems more valuable to report latencies rather than hypothetical velocities. In human subjects, these latencies can be measured using otoacoustic emissions or evoked potentials.

Otoacoustic emissions are small acoustic signals generated in the hair cells as they participate in the active filtering process of cochlea. The movements of the hair cells generate acoustic energy which is transmitted back along the basilar membrane and out through the middle ear to the ear canal. Several studies of human otoacoustic emissions have shown delays of approximately 8 ms between 10 kHz and 1 kHz (Kimberley et al., 1993; O'Mahoney and Kemp, 1995; Bowman et al., 1997). Although this delay is related to the travelling wave, its exact nature is not yet understood. Some investigators (e.g., Neely et al., 1988) have considered it to represent an acoustical 'round trip' (to the source of the emission and then back to the external ear), and have used one half of the delay to measure the 'one-way' travelling wave delay. As we shall cover later in Section 4, this view may be overly simplistic.

Recording auditory nerve action potentials or auditory brainstem responses can also provide a measure of the travelling wave delay. The auditory nerve responses contain an additional delay in the synapse between the inner hair cell and the afferent fiber. The auditory brainstem responses are affected by both this delay and a conduction delay between the auditory nerve and the generator of the response in the brainstem. Responses specific to particular regions of the cochlea can be obtained using clicks and derived band masking (e.g., Don and Eggermont, 1978) or brief tones and notched noise masking (e.g., Picton et al., 1979). Frequency-related delays in the responses from the cochlear nerve and auditory brainstem are similar to those estimated from one half of the delay of the otoacoustic emissions (Neely et al., 1988; Eggermont et al., 1996). This suggests that the synaptic and conduction delays occurring between the cochlea and the generator of the brainstem response are not significantly affected by the frequency of the sound.

Auditory steady-state responses offer several advantages over transient evoked potentials (Galambos et al., 1981; Stapells et al., 1984). Auditory steady-state responses evoked by the sinusoidal modulation of the amplitude of a tone contain energy only at the carrier frequency and sidebands separated from the carrier by the frequency of modulation. These responses are more frequency-specific than the responses to transient stimuli, which show a broad smearing of frequencies. Another advantage is that multiple steady-state responses can be evaluated simultaneously, provided each stimulus has a different modulating frequency. The separate responses are then measured in the frequency domain at the frequencies equal to the different modulating frequencies (Lins and Picton, 1995).

A steady-state response is characterized by its amplitude and phase. Even though the phase data are always available, most experiments have focused only on amplitude. The most frequent measurement of phase calculates 'phase coherence' (Fridman et al., 1984; Stapells et al., 1987; Dobie and Wilson, 1993; Rance et al., 1995): a response is considered reliable if its phase remains stable over time rather than varying randomly. Phase is not much used otherwise, probably because of ambiguities inherent in its measurements. The first ambiguity derives from the fact that phase is a circular rather than linear measurement (Fisher, 1993). When phase crosses an upper or lower limit (e.g., from 359° to 0°), the discontinuities in the measurement make normal linear calculations inappropriate. For example, the average of two measurements with phases of 359° and 1° is 0° rather than 180°. The second ambiguity is due to the steady-state nature of the response. It is impossible, from one measurement, to determine how many cycles occurred between the stimulus and the measured evoked steady-state response. A delay equivalent to one, two, or more cycles of the modulation frequency might have intervened between the stimulus and the response. One approach to this ambiguity is the measurement of 'apparent latency' (the slope of the phase versus frequency plot) introduced by Regan (1966). In the literature on auditory physiology, apparent latency is described as 'group delay' (Goldstein et al., 1971) and has been used to measure the latency of the motion of the basilar membrane (Patuzzi, 1996), the cochlear microphonic (Dallos and Cheatham, 1971), the otoacoustic emissions (Kimberley et al., 1993; Schneider et al., 1999), and the discharge of auditory nerve fibers (Anderson et al., 1971; Gummer and Johnstone, 1984; Joris and Yin, 1992).

This study examines the phases of auditory steady-state responses to discover if they relate to the acoustic frequencies of the eliciting stimuli in an orderly and stable manner. If these phases can be sensibly converted into latencies, and if these latencies are related to fre-

quency-related delays in the cochlea, then auditory steady-state responses might be useful in assessing travelling wave delays in normal and pathological ears. We used auditory steady-state responses evoked by tones modulated at rates greater than 75 Hz. Since these can be reliably recorded in infants (Rickards et al., 1994; Lins et al., 1996) and since they do not change significantly with sleep (Cohen et al., 1991; Lins and Picton, 1995), these rapid steady-state responses are clinically more useful than the slower 40 Hz responses.

2. Methods

2.1. Subjects

The 34 subjects (13 female) participating in these experiments were volunteers obtained from laboratory personnel, colleagues, and friends. Their ages ranged between 15 and 36 years. All subjects were screened for normal hearing for pure tones (range 500–6000 Hz) at 20 dB HL. During the recordings the subjects lay in a comfortable reclining chair inside a quiet room. Subjects were encouraged to relax and fall asleep during the recordings in order to reduce the background noise levels in the EEG. Most subjects slept for the entire recording period. Eight subjects participated in experiments 1 and 2. Six of the eight subjects in experiment 2 had also been subjects in experiment 1. Experiment 3 was performed concurrently with experiments 1 and 2 and used the same subjects. Experiment 4 (10 subjects) and experiment 5 (eight subjects) used data that had been recorded for a previous paper (John et al., 1998). Experiment 6 evaluated the replicability of the results in six of the eight subjects who had been studied in experiment 5. In experiment 7, we increased the number of subjects to 16 (eight female) since we wanted to explore male/female differences. These experiments were part of a larger research project on 'Evoked Potential Audiometry', which was approved by the Research Ethics Committee of the Baycrest Centre for Geriatric Care.

2.2. Auditory stimuli

The stimuli were generated digitally and converted to analog form with 12-bit resolution. The digital-to-analog (DA) conversion was based on a 4 MHz clock and occurred at a rate of once every 100 clock ticks (0.025 ms) or 40 kHz. The analog waveforms were then low-pass filtered (48 dB/octave) at 10 kHz to remove digitization noise and high-pass filtered at 250 Hz to remove any possible electrical artifact at the modulating frequencies. The stimuli were amplified to a calibration intensity and then attenuated to achieve the desired dB level. In experiments 4–7, the stimuli were presented to the subject through an Eartone™ 3A insert earphone inserted in the right ear. The stimuli were calibrated using a Brüel and Kjaer model 2230 sound level meter with a DB 0138 2 cm^2 coupler. For experiments 1–3, TDH-50P headphones were used and calibration was performed using a Brüel and Kjaer artificial ear type 4152. Both the ear inserts and the headphones show a relatively flat (± 3 dB) intensity response for stimuli of 4000 Hz and below. Stimuli of 6000 Hz had intensities that were 20 dB lower than the other stimuli for the ear insert and 5 dB lower for the headphone. (This large change in acoustic energy will be addressed in Section 4.)

Each amplitude-modulated (AM) stimulus was created by multiplying together two sine waves. The sine wave with the higher frequency (f_c) was the carrier; the sine wave with the lower frequency (f_m) formed the modulating envelope. The formula was:

$$a \times \sin(2\pi f_c t) \times (\sin(2\pi f_m t) + 1)/2$$

where a was the amplitude of the carrier. This resulted in 100% amplitude modulation of the carrier. The stimuli for multiple stimulation were formed by summing the individual AM stimuli (Lins et al., 1995). We have previously described these signals in both the time and frequency domains (John et al., 1998).

The frequencies of both the carrier and modulation signals were adjusted so that an integer number of cycles occurred within each recording section (usually lasting 409.6 ms). This allowed the individual sections to be linked together without acoustic artifact and to be interchangeable during artifact rejection (John and Picton, 1999). Thus, a carrier frequency of 1000 Hz was adjusted to 1000.98 Hz so that 410 complete cycles occurred within each section, and a modulation frequency of 81 Hz was adjusted to 80.566 Hz so that 33 cycles occurred within a section. For simplicity, carrier frequencies are henceforth expressed to the closest 10 Hz and the modulation frequencies with only one digit after the decimal point. Since we have shown that the amplitudes of the individual responses are not significantly affected when adjacent carrier frequencies are 1 octave apart (John et al., 1998), the experiments used tone combinations having octave separations such as 750, 1500, 3000, and 6000 Hz (750 octave series), or 500, 1000, 2000, and 4000 Hz (500 octave series).

2.3. Recording

Gold-plated recording electrodes were placed at the vertex and at the posterior midline of the neck just below the hair line (7–8 cm below the inion). The

ground electrode was placed on the right side of the neck. The skin beneath the electrodes was abraded to ensure that inter-electrode impedances were below 5 kΩ at 10 Hz. Electroencephalographic (EEG) signals were collected using a bandpass of 10–300 Hz (6 dB/octave).

The timing for the AD conversion was based on the same 4 MHz clock as controlled the DA conversion. Digitization of the EEG occurred every 3200 ticks of the clock (0.8 ms) at a rate of 1250 Hz (exactly 1/32 the rate of DA conversion).

Recordings continued over multiple sweeps each lasting 6.5536 s (8192 time points). Averaging was performed over 64 sweeps. Each recording sweep consisted of 16 sections each lasting 409.6 ms. In order to allow reasonable artifact rejection, a section (rather than an entire sweep) was rejected if the section contained any potentials with amplitudes greater than ± 40 μV. If a section was rejected, that part of the recording sweep was filled in with the next recorded data. Between 0 and 20% of the sections were rejected. Each recording lasted 7–8 min, depending on the number of rejections. Fig. 1 shows data recorded from a typical subject.

The preceding details describe the recording setup for all experiments except the last (experiment 7). This experiment used a new software program (John and Picton, 2000) with the following settings: an AD conversion rate of 500 Hz, a section size of 2048 points, and four sections per sweep yielding a sweep duration of 16.384 s. The modulation frequencies were chosen so that the numbers of cycles within a section (4.096 s) were adjacent integers (e.g., 328, 329, 330 and 331 giving modulation frequencies of 80.078, 80.322, 80.566, and 80.811 Hz).

2.4. Amplitude and phase measurements

Responses were initially analyzed by averaging together the recorded sweeps in the time domain. The resultant 6.5536 s (8192 time-point) averaged sweep was then transformed into the frequency domain using a fast Fourier transform (FFT). The FFT converts the original amplitude-time waveform into a series of cosine waves with specific frequencies, each represented as a vector on a two-dimensional plane. The specific frequencies available from the FFT are integer multiples of the $1/(Nt)$ resolution of the FFT where N is the number of timepoints and t is the time per timepoint. For experiments 1–6, the resolution of the FFT spectrum was 0.1526 Hz, and the resultant spectrum spanned from zero to 4096 times this amount or 625 Hz. (For experiment 7 the resolution was 0.061 Hz.)

For each frequency represented in the FFT spectrum, the X-Y (or real and imaginary) coordinates are then transformed into a vector with an amplitude and phase. The amplitude is the length of the vector ($\sqrt{X^2 + Y^2}$) and the phase is the rotation of the vector in relation to the X axis as calculated by arctan (Y/X). All phase measurements were expressed in terms of degrees rather than radians. When combining values across subjects, vector averaging procedures were used: each data point was represented as X and Y, averaging was performed separately for the X and Y values, and the average X and Y values then used to calculate the average amplitude and phase. In order to ensure that each subject contributed equally to the phase measurements, the data were normalized to unit amplitude before vector averaging.

2.5. Signal-to-noise assessments

The presence or absence of a response was assessed using an F ratio. This method evaluates if the response at the frequency of stimulation is different from the noise as estimated at adjacent frequencies where there is no response (Zurek, 1992; Dobie and Wilson, 1996; Lins et al., 1996; Valdes et al., 1997). The F statistic computes the ratio between the power measured at the signal frequency and the average power in the neighboring frequency bins. We estimated the noise levels using 120 frequency measurements – 60 above and 60 below the signal frequency (i.e., about 10 Hz on each side of the signal). When using the multiple stimulus technique those frequencies at which another signal is present were excluded from this calculation. The significance of this F ratio was evaluated against critical values for F at 2 and 240 degrees of freedom. For the data reported in this paper, 91% of the responses reached the $P < 0.001$ level of significance and only 4% did not reach the $P < 0.05$ level. Most of the responses that did not reach significance were evoked by the low carrier frequencies (500 or 750 Hz). Data which did not reach significance were still included in the combined measurements, since we were interested in how well the results represented normal findings (noisy though they might be).

2.6. Experimental design

2.6.1. Experiment 1. Monaural stimulation 80–100 Hz

In this experiment we used a 750 Hz octave series (i.e., 750, 1500, 3000 and 6000 Hz) of carrier frequencies and four different modulation frequencies in the 80–100 Hz range (80.6, 85.5, 90.3 and 95.2 Hz). The four different carrier frequencies were presented simultaneously with each being modulated by a different modulation frequency. Four separate recording conditions allowed us to measure the responses to each carrier frequency as modulated by each of the modulating frequencies.

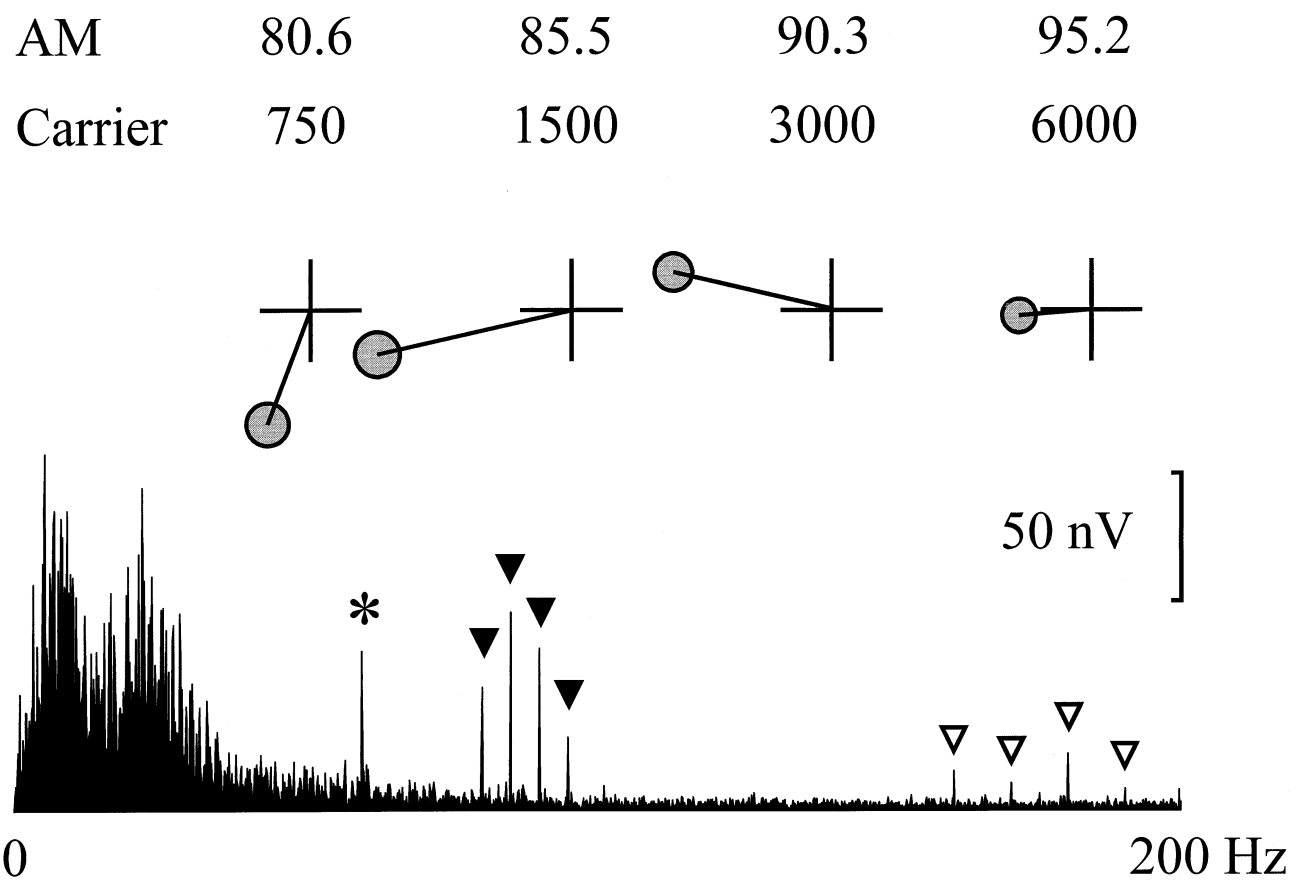

Fig. 1. Human auditory steady-state responses. This figure presents the data from a single subject in one of the conditions of experiment 1. Four tones were simultaneously presented. Each tone had its own specific carrier frequency and modulation frequency as shown at the top of the figure. The bottom section of the figure displays the amplitude spectrum of the averaged response up to 200 Hz. EEG alpha band activity and potentials related to movement and eyeblinks appear in the lower spectral frequency region. The peak indicated with the asterisk represents line noise at 60 Hz. The four peaks indicated with the filled triangles occur at the modulation frequencies of the stimuli and represent the responses specific to each of the four simultaneously presented stimuli. The unfilled triangles point to responses that represent the second harmonics of these responses, which were particularly evident in this subject. Polar plots of the responses at the modulation frequencies are shown in the center of the figure. The circles represent the $P < 0.05$ confidence limits of the noise in the adjacent frequency bins. For this figure the phase angles of the polar plots have been rotated to compensate for the phase delays caused by the amplifiers and the 90° shift caused by the stimulus being coded sinusoidally and the onset phase of the response being measured as a cosine.

2.6.2. Experiment 2. Monaural stimulation 150–190 Hz

This experiment was the same as experiment 1 except modulation frequencies in the 150–190 Hz range (158.7, 168.5, 178.2, and 187.9 Hz) were used. Six of the eight subjects were the same as in experiment 1.

2.6.3. Experiment 3. Dichotic stimulation

Experiments 1 and 2 each had two additional conditions in which stimuli were presented dichotically. In the first dichotic testing condition a 500 octave series was presented to the left ear using modulation frequencies of 78.1, 83.0, 87.9, and 92.8 Hz (153.8, 163.6, 173.34, and 183.11 Hz in experiment 2) while a 750 octave series was presented to the right ear using modulation frequencies of 80.5, 85.4, 90.3, and 95.2 Hz (158.7, 168.5, 178.2, and 187.9 Hz in experiment 2). In the second dichotic testing condition, these stimuli were presented to the opposite ears. Hence, in each dichotic condition, eight stimuli were simultaneously presented to a subject.

2.6.4. Experiment 4. Effects of intensity

Since latency increases with decreasing intensity with transient evoked potentials, we examined the effects of intensity on phase using a 500 octave series modulated in the 80–100 Hz range which was presented at both 35 and 75 dB SPL. This experiment was described in an earlier paper (John et al., 1998) although the phase data were not examined.

2.6.5. Experiment 5. Effects of multiple stimuli

This experiment examined the changes in latency that might be caused by presenting several tones simultaneously, rather than separately, as has been previously done. Responses to a single 1000 Hz tone or a single 2000 Hz tone modulated in the 80–100 Hz range served

as baseline measurements. These tones were then simultaneously presented with other tones in several conditions: two tones (the 1000 Hz and 2000 Hz tones together), four tones in one ear separated by one octave, four tones in one ear separated by one half octave, and eight tones in one ear separated by one half octave. Similar to experiment 4, the amplitude data from this experiment have been reported elsewhere (John et al., 1998).

2.6.6. Experiment 6. Stability over time

In order to assess the stability of steady-state phase, six subjects were recorded at two separate times 3–4 weeks apart using the 750 Hz octave series at both the 80–100 Hz and 150–190 Hz modulation frequencies. (The first recording session was the same as that in experiments 1 and 2.)

2.6.7. Experiment 7. Small differences in modulation frequency

In experiments 1–6 we attempted to compensate for using different modulation frequencies for each carrier. The effects of modulation frequency can be almost minimized by a method which uses almost identical modulation frequencies for all the carriers being tested. In this experiment we used modulation frequencies that were separated by only 0.244 Hz rather than the approximately 5 Hz used in the previous experiments. This was accomplished by increasing the sweep length to obtain 0.061 Hz resolution in the amplitude spectrum. The increased resolution caused adjacent modulation frequencies to be separated by three other bins in the amplitude spectrum. A baseline, using only the 1500 Hz carrier frequency (modulated at 80.322 or 160.156 Hz), was recorded in order to compare this response to that obtained using multiple stimuli with very small differences between adjacent modulation frequencies. Condition 1 of this experiment used modulation frequencies in the 80–100 Hz range (750 Hz modulated at 80.078, 1500 Hz at 80.322 Hz, 3000 Hz at 80.566, and 6000 Hz at 80.811). Condition 2 used modulation frequencies in the 150–190 Hz range (159.912, 160.156, 160.400, 160.645 Hz). Eight male and eight female subjects participated in this experiment.

2.7. Statistical evaluations

The main analyses concerned changes in the latency or phase of the responses which occurred due to the experimental manipulations. Prior to analysis, the measured data were arranged so that the rows of the data set corresponded to a carrier frequency and the columns corresponded to modulation rate. The statistical significance of the results was assessed by repeated-measures analyses of variance (ANOVA) with Greenhouse-Geisser corrections when appropriate (SPSS version 7.5). Because many of the post-hoc comparisons occurring in a single experiment were independent, we chose to use the Fisher least significant difference (LSD) test (GB-STAT version 5.0). Differences were considered significant at the $P < 0.05$ level.

2.8. Converting phase to latency

The recorded phase measurements were adjusted in several steps in order to remove delays introduced into the recordings by the equipment that generated the stimuli and recorded the responses. First, there were small delays caused by the 250 Hz high-pass filter that was used to prevent possible contamination of the stimulus with energy at the modulation frequencies. This delay varied with modulation frequency. Second, there were slight changes (varying with carrier frequency) in the timing of the stimuli related to electromechanical filtering in the insert microphone. Third, there was an acoustic delay of 0.9 ms in the tube connecting the speaker with the ear insert. These time delays were originally measured in milliseconds, using the acoustic calibration equipment, and converted into phase by multiplying by 360 and then dividing by the period of the modulation frequency. Fourth, the recording amplifiers caused changes in phase which could be measured by recording the response of the amplifier to a calibration signal at the modulation frequency. Fifth, the measured phase was then adjusted by 90° since the amplitude spectrum was computed in terms of cosine phase, whereas the stimulus was modulated using a sine function.

These final phase values (θ) were converted into phase delay (Rodriguez et al., 1986) using the equation:

$$P = 360 - \theta \quad (1)$$

Phase delay is more easily comprehensible than the onset phase, since it gets larger as the time of the response moves further away from the stimulus. The phase delays were maintained within the 0–360° limits by adding or subtracting 360 from measurements below or above this range, respectively.

Phase delay (P) can be converted to latency (L) according to the formula:

$$L = P/(360 \times f_m) \quad (2)$$

This is equivalent to multiplying the period of a full cycle ($1/f_m$) by the fraction of the cycle ($P/360$) represented by the phase delay. Since this does not take into account two ambiguities in the measurements, the for-

mula was adapted to:

$$L = (P + n\times 360 + m\times 360)/(360\times f_{\mathrm{m}}) \quad (3)$$

where n and m are integers selected to resolve the ambiguities associated with phase measurements.

The first ambiguity occurs because of the circularity of the phase measurement. When phase increases 1° from 360° to 1° the numerical measurement abruptly changes by 359°. This can be compensated for by 'unwrapping' the measurements (which have been bundled together in one cycle) and allowing the measurements to enter an adjacent cycle when they cross a cycle boundary. This ambiguity is illustrated diagrammatically in the upper section of Fig. 2. Fig. 3 shows the ambiguity in real data measuring the phase delays for a 750 Hz carrier modulated at different frequencies for eight subjects. The phase delay increases as the modulation frequency increases and, in the original phase values of four of the subjects, the phase jumped back to values close to zero (dashed lines). These values are not far apart from the other values (between 270° and 360°) in terms of circular phase. This ambiguity can be removed by adding 360° to those measurements that were more than 180° lower than adjacent measurements (in Eq. 3 n is set to 1). We are effectively proposing that a delay corresponding to a full cycle of the modulation frequency occurs before the measured responses in these subjects.

The second ambiguity arises from the steady-state nature of the response. Because the responses are the same from cycle to cycle, the number of cycles of the stimulus that might have occurred before the measured response is uncertain. The bottom section of Fig. 2 illustrates this ambiguity. This ambiguity can be resolved by measuring the response phase when the same carrier frequency (and, therefore, probably the same travelling wave latency) is modulated by several different modulation frequencies (each in their own recording session). This is illustrated by the data graphed in Fig. 4, which shows the effects of modulation frequency on the responses at different carrier frequencies. For a carrier frequency of 1500 Hz the phase delay was 152° at a modulation frequency of 80.6 Hz and 245° at 95.2 Hz (open circles in left graph of Fig. 4). If phase delays were directly converted into latencies according to Eq. 2 then latencies of 5.2 and 7.2 ms would be obtained since the period is 12.4 ms for an 80.6 Hz signal and 10.5 ms for a 95.2 Hz signal. The 2.0 ms difference is similar to the difference between the periods of the two modulation frequencies (12.4−10.5 = 1.9 ms). If we therefore propose that a delay equivalent to one full cycle occurred before the cycle being measured ($m = 1$ in Eq. 3) the latencies become 17.6 ms (12.4+5.2) and 17.7 ms (10.5+7.2). The measurements are closer than with no preceding cycles and closer than what would be obtained if we postulated two preceding cycles ($m = 2$), which would give latencies of 30.0 and 28.2 ms.

We based the choice of m on the goodness of fit of the collected data when multiple modulation frequencies were used with the different carrier frequencies. This is illustrated in Fig. 5. The left graph assumes that no preceding cycles have occurred before the response. There are significant effects of both carrier frequency and modulation frequency. The middle graph shows the results of adding a cycle ($m = 1$) to the data shown in the upper graph. In this graph, the different modulation rates yield approximately equivalent estimates of latency. As can be seen in the right graph, adding a further cycle ($m = 2$) causes the data to diverge again with a significant effect of modulation frequency. Once m has been decided on the basis of these results, m can be applied to responses recorded using only one modulation frequency for each carrier frequency (e.g., experiment 3). An ambiguity might arise if there is no clear difference in the effect of modulation frequency for two adjacent values of m.

This approach is related to the 'apparent latency' calculated by dividing the slope of phase delay versus stimulus frequency by 360 (Regan, 1966, 1989; see also van der Tweel and Verduyn Lunel, 1965). In this particular example, the apparent latency of the two responses to 1500 Hz carriers is (245−152)/(360×(95.2−80.6)) or 17.7 ms. A graphic way of calculating this measurement was proposed by Diamond (1977). The equivalence of these different measurements of latency is documented in the Appendix. As already stated, apparent latency is often referred to as 'group delay' (Goldstein et al., 1971) in the auditory literature.

Unfortunately, it is often difficult to interpret apparent latency since the phase shifts occurring in physiological filters are difficult to characterize (Bijl and Veringa, 1985) and the actual relation of apparent latency to physiological delay is not clear (Hari et al., 1989). A major problem occurs when the physiological response shows filter effects along the dimension of the modulation frequency. The physiological system may only respond to a band of modulation frequencies or may respond less to higher modulation frequencies. A first order low-pass filter will show a phase shift equal to $\arctan(f_{\mathrm{m}}/f_0)$ where f_0 is the cut-off frequency of the filter (Brown et al., 1982). It might be possible to estimate the effect of such filtering on the responses from the amplitudes of the responses at the different modulation frequencies, and then compensate the phase measurements prior to latency calculations (Regan, 1989, pp. 29–30). We performed these calculations for our data based on estimated phase shifts of about 90° per octave, but these did not significantly alter the results.

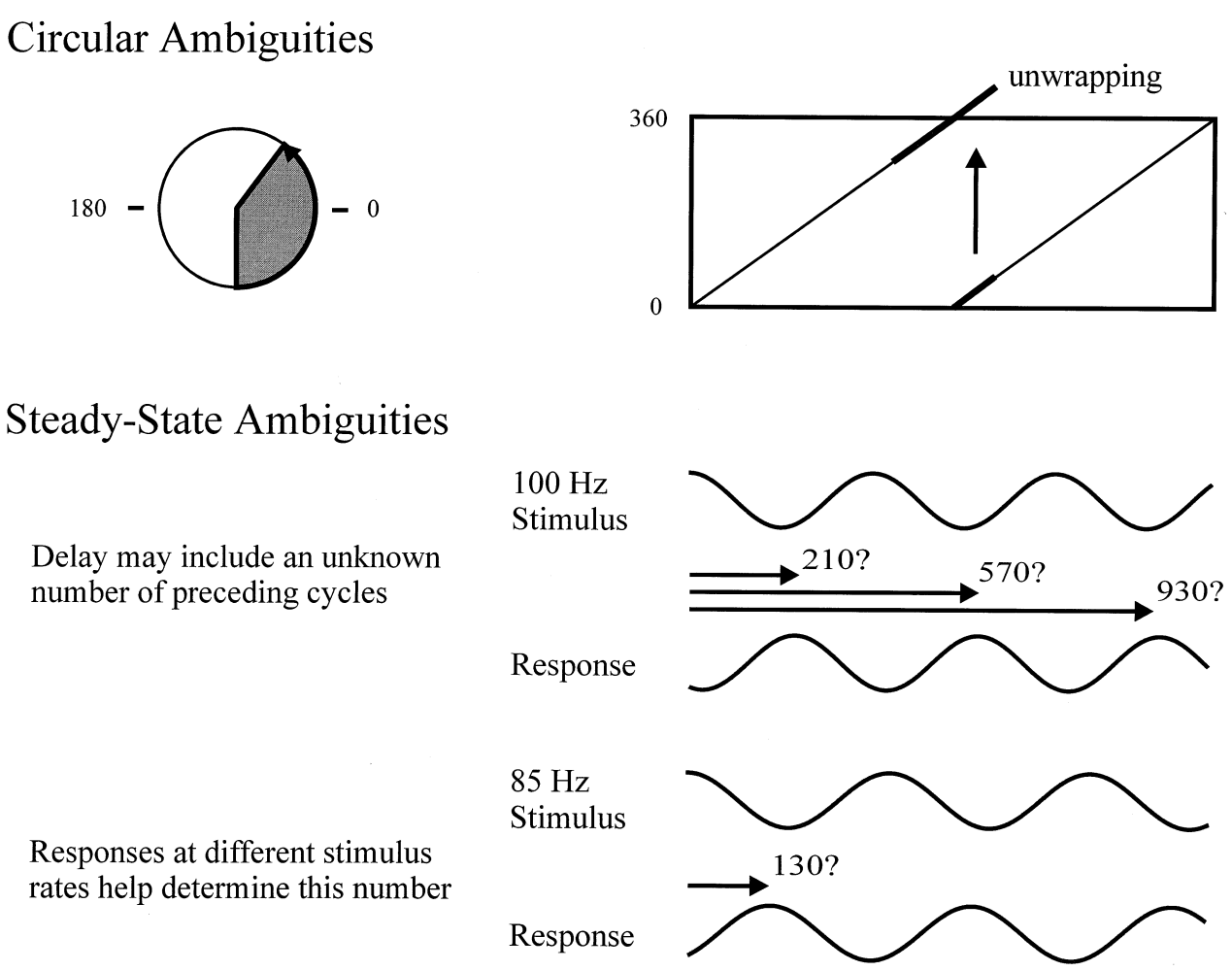

Fig. 2. Ambiguities in the measurement of phase. The upper half of the figure shows the ambiguities that occur because phase is a circular rather than linear measurement. If the phase increases as some other parameter is varied, it may exceed 360° and revert to 0°, causing a sudden fall in the regularly increasing measurements. This can be handled by 'unwrapping' the measurements by adding 360° (upgoing arrow in the upper right section of the figure). The bottom half of the figure illustrates the ambiguities that occur because the response is recorded in steady state. The cosine onset phase of the response recorded in the second line was measured as 150°. Phase delay is measured as the phase difference from the stimulus waveform (onset at 0° cosine) to the same part of the response waveform. This is equivalent to 360° minus the measured phase or 210°. Translated to latency, this is 210/(360×100) or 5.8 ms. Unfortunately, there is no way of knowing whether the 0° point of the ongoing stimulus is evoking the point on the response waveform where the phase is 210° or a point occurring one or two cycles later (at latencies 15.8 or 25.8 ms for 570° and 930° phase delays). If we record the response at another stimulus rate (and if we can assume that the response has the same latency for the two rates), then we can determine which of the possibilities is most likely. At 85 Hz the phase delay was 130°. This gives a latencies of 4.2, 16.0 and 27.8 ms. The measurements of 15.8 and 16.0 ms are the most similar between the two stimulus rates. An 'apparent latency' can also be calculated from the slope of the phase with stimulus frequency divided by 360. For these measurements, this latency is 80/(15×360) or 14.8 ms.

Another approach is to eliminate these effects by considering data (for different carrier frequencies) only at one modulation frequency. This can be done with multiple measurements (data from experiments 1 and 2), or by using modulation frequencies so close together that any effect of modulation frequency would be negligible (experiment 7).

2.9. Modeling

Some simple modeling was undertaken to determine how the measured latencies might occur on the basis of known physiological processes. A MATLAB program was constructed to imitate the effects of four processes. First, the filtering that occurs at the level of the hair cells was modeled as a gammatone filter using parameters from Patterson (1994; see also Glasberg and Moore, 1990). Second, the filtered signal was rectified. For simplicity and since we were not interested in the effects of different intensities, no compression (e.g., Corey and Hudspeth, 1983; Dallos, 1985) was used in the rectification. Third, we modeled the refractoriness of the auditory nerve fiber discharges using the equations of Carney (1993; see also Eggermont, 1985), and

parameters approximately equivalent to those recommended in that paper (an absolute refractory period of 1 ms and then a double exponential return with time constants of 1 and 25 ms). Since we were estimating population discharges rather than individual fiber discharges, the refractoriness was applied with a strength of 40%. This meant that only that portion of the neurons would fire at any one time and therefore be susceptible to being refractory the next time. Fourth, we estimated the compound field potentials generated from the population of neurons in the brain responsible for the scalp potentials. This involved imposing a latency jitter (arbitrarily set as a rectangular 2 ms function) to mimic the variability in the conduction times between the cochlea and the generator neurons. The field potential of a single fiber discharge was estimated using the 'model unit response' of Elberling (1976) (see also Teas et al., 1962) without the secondary peak. This was then convolved with the estimated population discharge pattern to give the modeled field potentials.

The latencies of the modeled responses were measured by determining the latency of the maximum of the cross-correlation function between the original acoustic signal and the modeled field potential. Two types of signals were analyzed: an AM tone with a modulation frequency of 80 Hz and carrier frequencies between 500 and 8000 Hz, and a click with a duration of 200 μs presented at a rate of 50/s.

3. Results

3.1. Experiment 1. Monaural stimulation 80–100 Hz

Sample data from one subject in this experiment are shown in Fig. 1. Fig. 3 shows the inter-subject variability of the phase delays at 750 Hz. The left side of Fig. 4 shows the mean phase delays for all four carrier frequencies and all modulation frequencies.

A 4×4 repeated-measures ANOVA for the phases of the responses to the 750 series modulated at 80–100 Hz, which compared carrier frequency and modulation frequency, indicated significant effects of both carrier frequency ($F = 33.8$; df 3,21; $P < 0.001$) and modulation rate ($F = 111.2$; df 3,21; $P < 0.001$), as well as a significant interaction ($F = 3.9$; df 9,63; $P < 0.05$, $\varepsilon = 0.31$) due to a flattening of slope as the carrier frequencies increased.

When converting from phase delay into latency, the data can be modeled as having occurred after 0, 1 or 2

Table 1
Estimated latencies at different modulation frequencies (data from experiments 1 and 2)

Rate (Hz)	Latency (ms)			
	750	1500	3000	6000
80.6	20.6	17.6	16.1	16.4
85.5	20.6	17.8	16.3	16.2
90.3	20.8	17.7	16.1	15.9
95.2	20.8	17.7	15.8	15.9
158.7	11.3	8.6	6.6	6.3
168.5	11.0	8.8	6.8	6.5
178.2	10.9	8.8	7.0	6.5
187.9	12.0	9.0	7.0	6.5

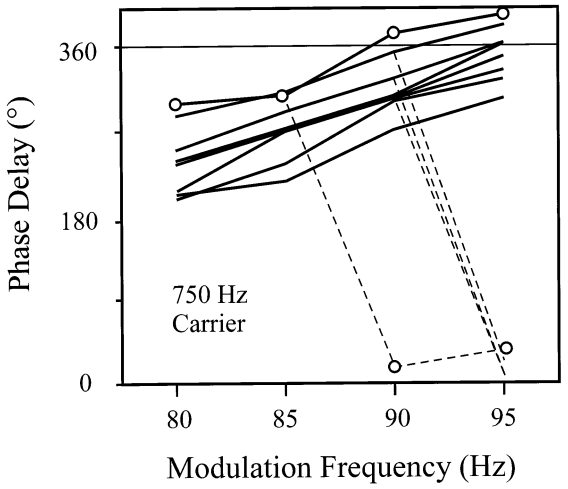

Fig. 3. Unwrapping of phase. This figure shows the phase of the response to the 750 Hz tone at different modulation frequencies in all eight subjects studied in experiment 1. The original measurements of phase delays were limited to the range 0–360°. In four of the subjects, this caused large discontinuities in the course of the lines (shown by dashes). These were compensated by adding 360° to the aberrant measurements. For example, in the subject whose measurements are identified using the circles the phase delay jumped from 310° at 85.5 Hz down to 14° at 90.3 Hz and then increased to 36° at 95.2 Hz. The latter two measurements were unwrapped to 374° and 396°. The figure also illustrates the inter-subject variability in the measurements.

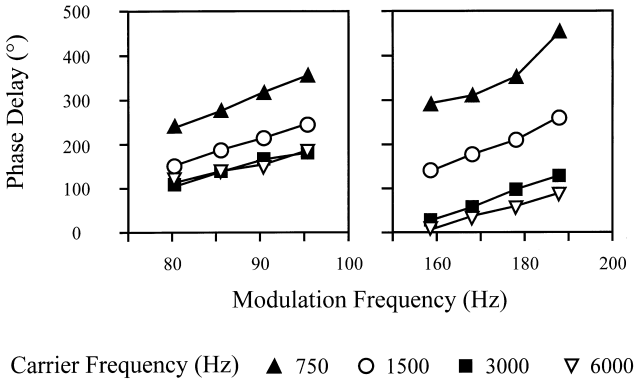

Fig. 4. Phase-frequency plots. This figure shows the average data from experiments 1 and 2. As can be seen in the graphs, the phase delay increases regularly with increasing modulation frequency with a slope that increases with decreasing carrier frequency. This slope can be used to estimate 'apparent latency' or 'group delay'.

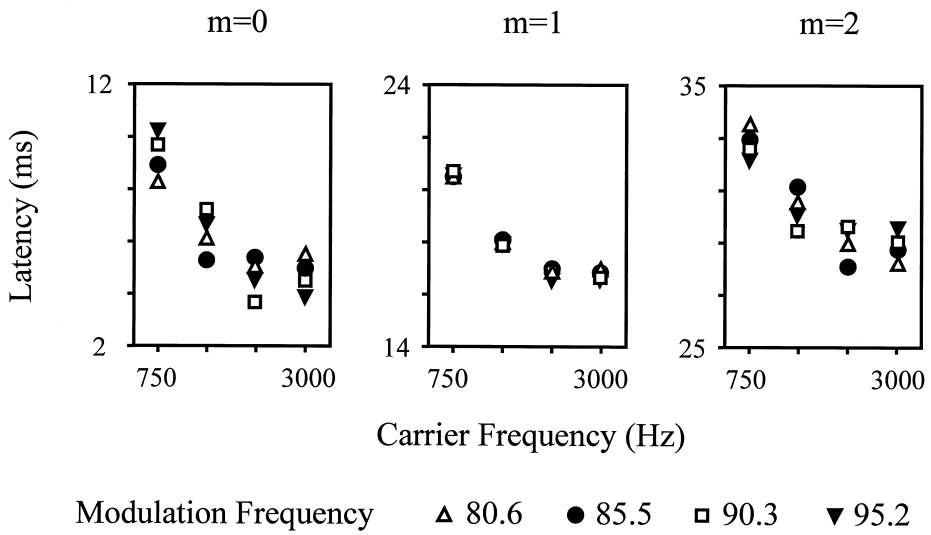

Fig. 5. Preceding cycles (80–100 Hz). This figure illustrates a method by which the number of cycles preceding the recorded response can be estimated. The phase delay is converted to latency and then the actual latency from the stimulus is calculated by adding an integer number (m) of cycles of modulation to the estimated latency. Since carrier frequencies are what should determine the delays in the cochlea, different modulation frequencies should result in the same latency estimate. As can be seen in the figure, the delays related to each carrier frequency are most consistent across the different experimental conditions when $m=1$.

preceding cycles by altering the parameter m in Eq. 3. These calculations are shown in Fig. 5. When $m=0$, the experimental analysis of the latency data showed main effects of both carrier frequency ($F=152.8$; df 3,21; $P<0.001$, $\varepsilon=0.43$) and modulation rate ($F=18.3$; df 3,21; $P<0.001$, $\varepsilon=0.42$). When $m=1$, there was a significant change in latency with carrier frequency ($F=101.2$; df 3,21; $P<0.001$, $\varepsilon=0.43$) but not with modulation rate ($F=0.56$; df 3,21; $P=0.7$). By adding a further cycle, so that $m=2$, the data again became divergent, and a significant effect of modulation rate again emerged ($F=33.8$; df 3,21; $P<.001$, $\varepsilon=0.43$).

The best fit for the data, when $m=1$, yields steady-state response latencies for the 750 octave series of 20.7, 17.7, 16.1, and 16.1 ms. Post-hoc Fisher LSD tests showed that the differences in latency between carrier frequencies were significantly different ($P<0.01$) except between 3000 Hz and 6000 Hz. Table 1 provides the latencies estimated for each carrier frequency at each modulation rate when $m=1$. The apparent latencies for these data were 22.2, 17.5, 13.9, and 12.7 ms.

The amplitudes are plotted in Fig. 6. There was no significant effect of carrier frequency ($F=3.57$; df 3,21; $P=0.058$, $\varepsilon=0.64$) or of modulation frequency ($F=2.54$; df 3,21; $P=0.13$, $\varepsilon=0.50$). The 750 and 6000 Hz tones tended to produce smaller responses. Furthermore, the overall amplitudes across the different carrier frequencies did show a slight decrease in amplitude with increasing modulation frequency (5 nV from 80 to 95 Hz) that was equivalent to about 6 dB/octave. Assuming that this may have been associated with a phase shift of 90° per octave, the phase measurements were adjusted and latencies recalculated. The differences indicated by these calculations were less than 0.2 ms.

3.2. Experiment 2. Monaural stimulation 150–190 Hz

The mean phase measurements are shown on the right of Fig. 4. The latencies are plotted with different values of m in Fig. 7. In the case of the 150–190 Hz series, for $m=0$, a 4×4 repeated-measures ANOVA showed a main effect of latency for carrier frequency ($F=165.5$; df 3,21; $P<0.001$, $\varepsilon=0.48$) and also for modulation rate ($F=59.6$; df 3,21; $P<0.001$, $\varepsilon=0.48$). When $m=1$, there was a significant change

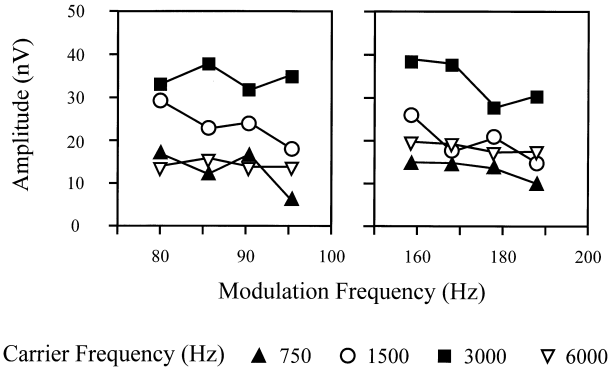

Fig. 6. Effects of modulation frequency on amplitude. This figure represents the mean data for the eight subjects evaluated in experiments 1 and 2. There is a slight decrease in amplitude with increasing modulation frequency. This was not significant between 80–100 Hz but was significant between 150–190 Hz. This suggests a small low-pass filtering process.

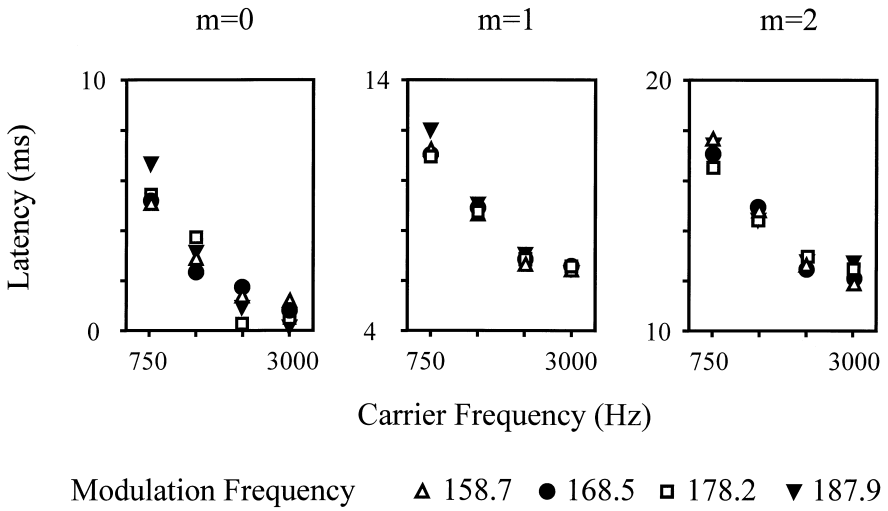

Fig. 7. Preceding cycles (150–190 Hz). This figure shows how the number of preceding cycles of modulation was estimated for the data recorded with modulation frequencies 150–190 Hz. The figure is similar to Fig. 4. Again, the data are most consistent across the recording condition when the number (m) of preceding cycles of modulation equals one.

in latency with carrier frequency ($F=96.4$; df 3,21; $P<0.001$, $\varepsilon=0.48$), but not for modulation rate ($F=1.2$; df 3,21; $P=0.3$). After adding a further cycle so that $m=2$, the data again became divergent across both carrier frequency ($F=165.5$; df 3,21; $P<0.001$) and modulation frequency ($F=8.9$; df 3,21; $P<.01$).

The mean latencies for the 750 series 150–190 Hz modulated data were 11.3, 8.9, 6.9, and 6.4 ms. Post-hoc Fisher LSD tests within each modulation frequency showed significantly ($P<0.01$) different latencies across the different carrier frequencies except for the difference between 3000 and 6000 Hz. However, all the latencies were significantly different when the data were collapsed across all modulation rates ($F=6.9$; df 1,14; $P<0.001$). The apparent latencies for these data were 15.1, 10.9, 9.5, and 7.5 ms.

The mean amplitudes of these responses are plotted in the right half of Fig. 6. The trends seen in the 80–100 Hz data again appeared here, but now there was a significant effect of carrier frequency ($F=8.04$; df 3,21; $P<0.05$) and also modulation rate ($F=6.74$; df 3,21; $P<0.05$). As in the case of the 80–100 Hz data, there again was a decrease in the amplitudes of the 750 and 6000 Hz responses and a roughly linear decrease in amplitude as modulation rate increased (7 nV from 155 to 185 Hz). Again, recalculation of the latencies after compensating the phases for a possible 90°/octave shift did not significantly alter the latencies.

3.3. Experiment 3. Dichotic stimulation

In the first dichotic condition, a 750 Hz octave series was presented to the right ear while a 500 Hz octave series was presented to the left ear, and in the second condition these stimuli were presented to the opposite ears. In order to analyze the data, the results were combined to provide latency estimates at one half octave steps in each ear. The mean data compared between the ears are plotted in Fig. 8. The means and standard deviations across the subjects are provided in Table 2. There are minor differences between Fig. 7 and the Table 2 data since vector averaging was used to calculate the mean phases for the latencies in the figure but ordinary arithmetic averaging was used to calculate the latencies for the table (to allow us to obtain measurements of inter-subject variability).

The responses to the 80–100 Hz dichotic AM tones showed significant effects for carrier frequency ($F=71.2$; df 7,49; $P<0.001$) and no significant effect of ear ($F=2.15$; df 1,7; $P=0.19$). Post-hoc Fisher LSDs

Table 2
Effects of carrier frequency on response latency (data from experiment 3)

Modulation range (Hz)	Carrier(Hz)							
		500	750	1000	1500	2000	3000	4000	6000
80–100 Hz	mean	21.5	20.2	18.9	17.6	16.7	16.0	16.1	15.8
	S.D.	1.6	1.5	0.8	0.6	0.5	0.5	0.7	0.8
150–190 Hz	mean	12.2	11.1	10.3	8.9	7.8	6.8	6.9	6.6
	S.D.	1.7	1.0	1.1	1.4	0.7	0.3	0.4	0.5

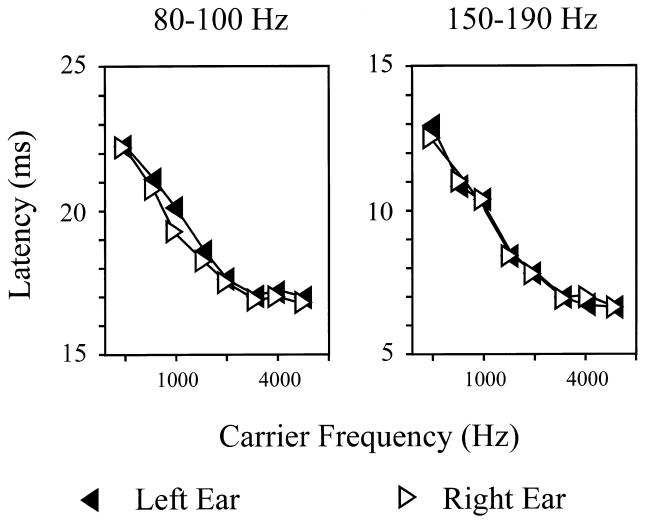

Fig. 8. Dichotic stimulation. This figure shows the results of experiment 3. Four stimuli with carrier frequencies separated by an octave were presented to one ear and four stimuli at intervening carrier frequencies were presented to the other ear. The data are the latencies derived using $m=1$ from the vector-averaged phase delays across the eight subjects recorded in this experiment.

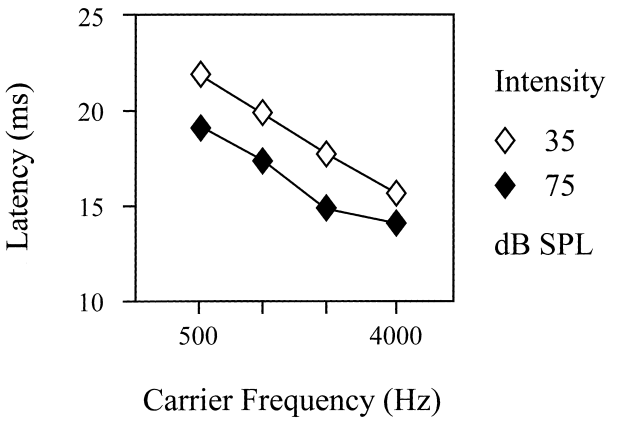

Fig. 9. Effects of intensity on response latency. These data represent the average results across the eight subjects participating in experiment 4. They show a decrease (approximately 2 ms) in the estimated latency with increasing intensity.

indicated that the differences in latency between all adjacent carrier frequencies (1/2 octave steps) were significant at the 0.001 level up to the 3000 Hz level. The responses to the 150–190 Hz dichotic AM tones showed significant effects for carrier frequency ($F=91.8$; df 7,49; $P<0.001$) and, again, the responses for the left and right ears failed to show any significant differences ($F=0.07$; df 1,7; $P=0.8$). Post-hoc Fisher LSDs again indicated that the differences in latency between all adjacent carrier frequencies (1/2 octave steps) were significant at the 0.001 level up to the 3000 Hz level. We also compared the responses recorded when stimuli were presented to the left ear in the monaural condition to those recorded in the left ear during the dichotic condition. There were no significant differences between monaural and dichotic response latencies.

3.4. Experiment 4. Effects of intensity

When the intensity of the stimuli increased from 35 dB SPL to 75 dB SPL (Fig. 9), the latency of the response in the 80–100 Hz range decreased by an average of 2.4 ms (2.6, 2.5, 2.8, and 1.7 ms for the 500, 1000, 2000, 4000 Hz carrier frequencies). A comparison between 35 and 75 dB SPL response indicated significant changes with intensity ($F=30.9$; df 1,7; $P<0.001$) and with carrier frequency ($F=33.6$; df 3,21; $P<0.001$). Although there was a smaller decrease at the 4 kHz response there were no significant interactions between loudness and carrier frequency with respect to latency. Post-hoc LSD comparisons indicated that this decrease was significant for all frequencies at the 0.05 level.

3.5. Experiment 5. Effects of multiple stimuli

In the other experiments presented in this paper four stimuli were presented in each ear. Here, we examined the latencies obtained by a 1000 Hz or 2000 Hz tone presented alone, compared to the latencies obtained when multiple stimuli were used. The latencies of the 1000 Hz and 2000 Hz tones in the different stimulus conditions are shown in Table 3. A two-way (carrier frequency by number of simultaneously presented stimuli) repeated-measures ANOVA showed a significant main effect for carrier frequency ($F=103.2$, df 1,9; $P<0.001$) but not for number of stimuli ($F=2.13$; df 4,36; $P=0.10$). However, there was a significant interaction between carrier frequency and number of stimuli ($F=7.96$; df 4,36; $P<0.001$). Compared to the single stimulus condition, when adjacent stimuli were separated by less than 1 octave, increased latency was seen, although post-hoc testing showed this effect was

Table 3
Effects of multiple stimuli on latency estimates (data from experiment 4: means and standard deviations)

Carrier (Hz)	One tone	Two tones (1 octave)	Four tones (1 octave)	Four tones (1/2 octave)	Eight tones (1/2 octave)
1000	20.9 ± 2.9	21.5 ± 2.8	21.3 ± 2.7	22.4 ± 2.8	21.9 ± 2.9
2000	18.9 ± 2.7	18.3 ± 2.5	18.7 ± 2.7	19.1 ± 2.9	19.9 ± 2.8

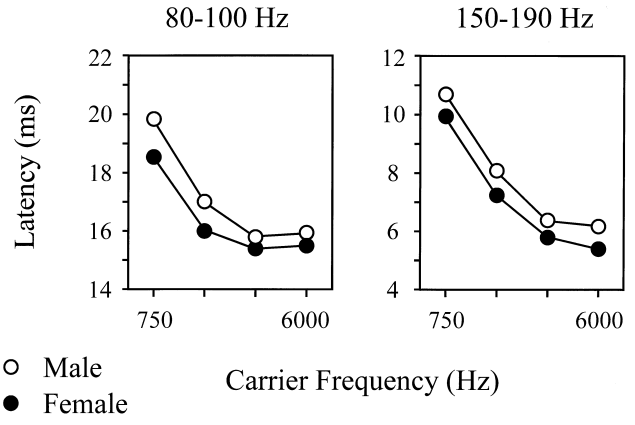

Fig. 10. Small differences in modulation frequency. This figure presents the results of experiment 7. The latencies were slightly shorter for the eight female subjects than for the eight male subjects.

only significant ($P < 0.05$) in the case of the 1000 Hz tone.

3.6. Experiment 6. Stability of phase

In order to assess the stability of steady-state phase we tested six subjects and then tested them again 3–4 weeks later, at both 80–100 Hz and 150–190 Hz modulation rates. In the 80–100 Hz condition the average change was −0.11 ms (range: −0.66 to +0.67 ms). In the 150–190 Hz condition the average change was only −0.04 ms (range: −0.49 to +0.31 ms). The results for each carrier frequency are provided in Table 4. For the 80–100 Hz responses, the latency difference between 750 and 6000 Hz changed from one recording to the next by 0.15 ms (range: −0.71 to +1.14 ms). For the 150–190 Hz responses, this measurement changed by −0.06 ms (range: −0.25 to +0.05 ms).

3.7. Experiment 7. Small differences in modulation frequency

The latencies for this experiment are plotted in Fig. 10. The differences between the latency for the 750 Hz carrier and the latency for the 6000 Hz carrier were 3.5 and 4.5 ms for the 80 Hz and 160 Hz series, respectively. Like experiments 1, 2, and 3 the 80 Hz data failed to show any decrease in latency between the 3000 to 6000 Hz responses, while in the 160 Hz data, a decrease was again apparent.

An ANOVA of all of the data indicated highly significant effects for carrier frequency ($F = 107.7$; df 3,42; $P < 0.001$) and between the 80 and 160 Hz modulation rates ($F = 2840.0$; df 1,14; $P < 0.001$). There was also a significant effect of gender ($F = 4.6$; df 1,14; $P = 0.05$) with the male subjects showing latencies which were delayed, on average, by 0.78 ms compared to the female subjects. There was no significant interaction between gender and carrier frequency although the difference between 750 and 6000 Hz was slightly greater for males than for female subjects in the 80 Hz data.

In order to see the effect of such closely spaced modulation frequencies on the measurements, the amplitude and phase of the 1500 Hz tone presented alone was compared to its amplitude when it was presented simultaneously with the three other tones. For the 80 Hz response the amplitudes were smaller (92%) when the stimulus was presented with the other tones, while in the case of the 160 Hz response the amplitude was larger (140%). For both modulation frequencies the responses were slightly earlier by 0.4 ms (same for both) when the stimulus was presented with the other tones.

Table 4
Stability of the responses over two separate recording sessions (data from experiment 6)

Modulation	Carrier	Latency		Difference between sessions			
		mean	S.D.	mean	S.D.	min	max
80–100 Hz	750	19.9	0.8	0.1	0.4	−0.5	0.5
	1500	17.0	0.6	0.1	0.4	−0.3	0.7
	3000	15.4	0.9	0.2	0.4	−0.4	0.6
	6000	15.5	0.6	0.0	0.6	−0.7	0.7
150–190 Hz	750	10.8	0.8	0.0	0.1	−0.2	0.1
	1500	8.8	0.8	0.0	0.2	−0.3	0.3
	3000	6.3	0.4	0.0	0.2	−0.3	0.2
	6000	6.1	0.4	−0.1	0.2	−0.5	0.1

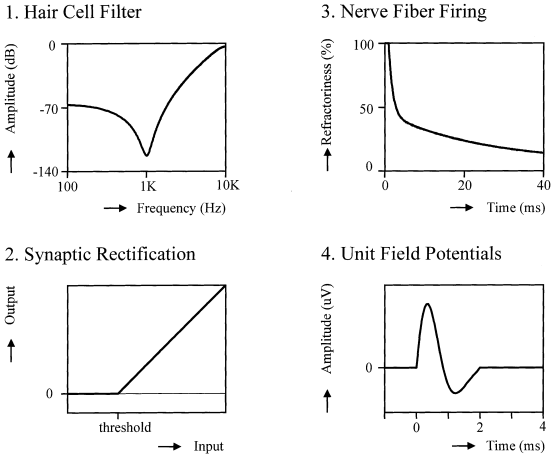

Fig. 11. Modeling parameters. This figure shows the parameters used in the four processes for modeling the latencies of the responses to amplitude-modulated tones and clicks. The first process is the cochlear filter which is modeled as a gammatone filter using the parameters of Patterson (1994). The second process is synaptic rectification. The third process is to estimate the fiber discharge rates. For this process, the refractoriness of the fibers was estimated using parameters from Carney (1993). The fourth process in the model was to estimate the field potentials generated from the responding units. As well as a jittering of the latency (not illustrated) the response pattern was then convolved with an estimated unit field potential using parameters similar to those of Elberling (1976).

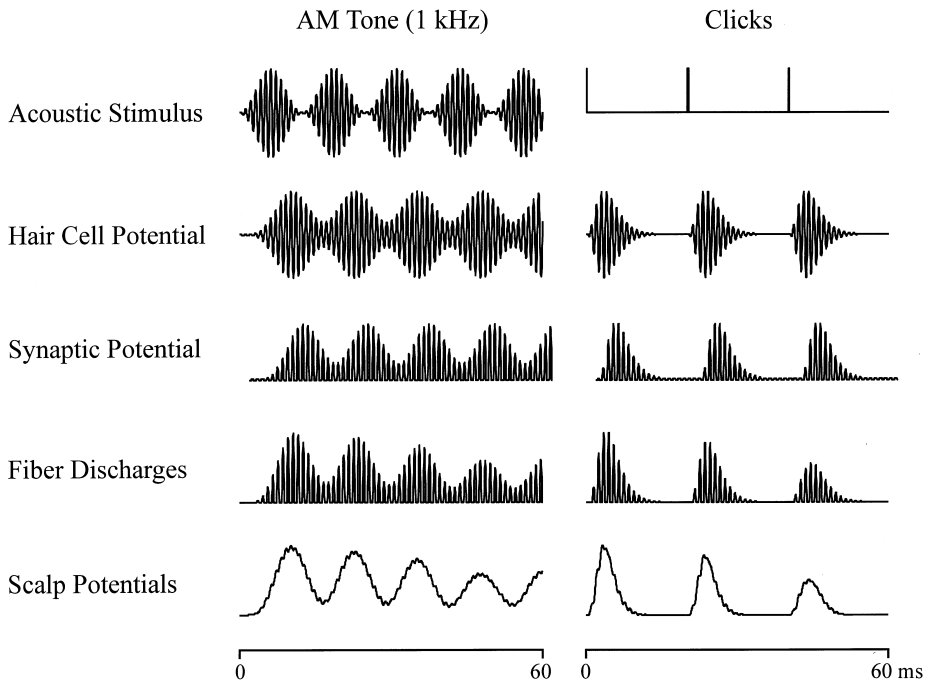

Fig. 12. Modeling the physiological processes. This figure illustrates the processing of acoustic stimuli through the various steps in the model. The hair cell filter intervenes between the acoustic stimulus and the hair cell potential. Rectification then occurs to give the synaptic potential. The refractoriness of the nerve fibers determines the fiber discharge pattern. The scalp potentials are then determined by both the jitter and the field pattern of the unit discharges.

3.8. Results of modeling

The basic parameters of the model are illustrated in Fig. 11. Fig. 12 shows how the different physiological components of the model might process two different inputs – an 80 Hz AM tone and a click presented at a rate of 50 Hz. Fig. 13 graphs the latencies estimated using cross-correlation for these two kinds of stimuli. The carrier frequency of the AM tone and the characteristic frequency of the hair cell filter were varied in octave steps from 500 to 8000 Hz. The latency changes over the complete range were 5.4 ms for the AM tone and 3.2 ms for the click-evoked potential. Estimates of the difference between 750 and 6000 Hz were 3.7 ms and 2.4 ms.

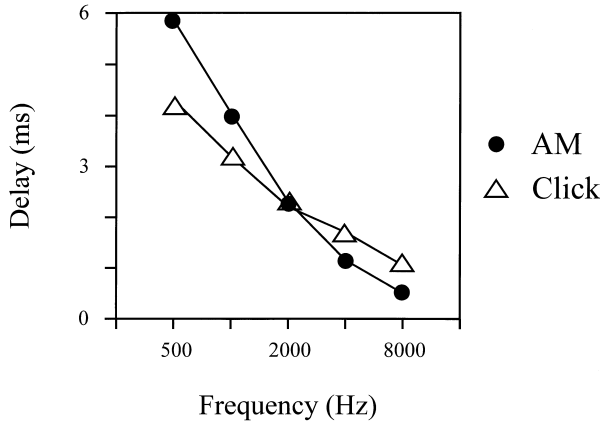

Fig. 13. Modeled latencies. This figures graphs the latencies obtained for the steady-state response to a pure tone amplitude-modulated (AM) at 80 Hz and a click stimulus presented at a rate of 50 Hz. The latency change of the steady-state response with changing carrier frequency is almost double that of the click response.

4. Discussion

4.1. Origins of the human auditory steady-state responses

Human auditory steady-state responses can be recorded using stimulus rates up to several hundred Hz (Rickards and Clark, 1984). The amplitude of these responses is larger at rates near 40 Hz (Galambos et al., 1981) and 80 Hz (Cohen et al., 1991; Lins et al., 1995). Responses at both these rates have been used as an objective measurement of hearing thresholds (Rance et al., 1995; Lins et al., 1996; Dobie and Wilson, 1998). The more rapid responses are less affected by changes in sleep or maturation (Cohen et al., 1991; Levi et al., 1993; Lins and Picton, 1995) and might therefore provide more reliable measurements in children and other difficult-to-test patients. The results provide in this paper indicate that responses at rates of 150–190 Hz might also provide reliable measurements.

The nature of human auditory steady-state responses

has been mainly studied with the responses evoked at stimulus rates near 40 Hz (Galambos et al., 1981; Stapells et al., 1984). If they largely arise from the same neurons that generate transient evoked potentials, the steady-state responses can be predicted from the superposition of multiple overlapping transient responses (Galambos et al., 1981; Hari et al., 1989; Plourde et al., 1991). However, it is also possible that the rapid stimulation might elicit a separate response from neurons that are specifically responsible for rhythmic activity and that resonate at the frequency of stimulation (Basar et al., 1987). Such generators might also be responsible for spontaneous cortical gamma rhythms. Discrepancies between the steady-state response and the response predicted by superposition (Azzena et al., 1995) and the persistence of the response beyond the time when the stimuli are presented (Santarelli et al., 1995) support the idea of resonance. However, by modeling the refractory effects and considering multiple generators, one can still explain most of the 40 Hz response findings on the basis of the transient response waveforms (Gutschalk et al., 1999). We therefore think it unlikely that that independent rhythmic generators might account for the 40 Hz steady-state responses. For the responses evoked by stimuli modulated at frequencies above 80 Hz, it is even more unlikely. We therefore consider these responses to be generated by neurons in the brainstem (and possibly cortex) that both respond to transient stimuli and become locked to the envelopes of amplitude-modulated tones.

The estimated latencies for the 150–190 Hz responses were about half the latencies for the 80–100 Hz responses (e.g., Table 4, Fig. 8). Our latencies (17.7 and 8.9 ms at 1500 Hz) are similar to those of Cohen et al. (1991) who reported apparent latencies at 1000 Hz (55 dB HL) of 24.8 ms for modulation frequencies of 30–60 Hz and 13.0 ms for 90–125 Hz, given that the latencies decrease with increasing modulation frequencies (Rickards and Clark, 1984). Previous work in our laboratory had provided an apparent latency of 19 ms for 1000 Hz tones modulated between 71 and 100 Hz. The latencies at 150–190 Hz suggest that the generators might be at about the same level of the brainstem as those generating wave V of the auditory brainstem response. The longer latencies for the 80–100 Hz responses suggest either a generator further along the auditory pathway or a generator lower in the pathway activated by a multisynaptic circuit.

4.2. Latency of steady-state responses

The phase of a steady-state response is related to delay. Unfortunately, this relationship is ambiguous and phase cannot be directly translated into latency. By assessing phase at different stimulus rates, it is often possible to resolve the ambiguities and arrive at a clear estimate of 'apparent latency' (Regan, 1966) or 'group delay' (Goldstein et al., 1971).

The distorting effect of physiological filters operating across the different rates of stimulation remains a problem (Bijl and Veringa, 1985). This is analogous to determining the delays related to filter slopes when studying the group delays of tuning curve (Goldstein et al., 1971). The effects of such filters can only be estimated. In the data presented here, these effects appear small since similar results were obtained when the latency estimates were based on the same modulation frequencies rather than different modulation frequencies. Supporting data were obtained on separate recordings for the different carrier frequencies with each carrier frequency being modulated by exactly the same modulation frequency (experiments 1 and 2 – Table 1), and using modulation frequencies that were sufficiently close together during multiple stimulation that any filter effect would be negligible (experiment 7).

The physiological interpretation of the latencies of the scalp-recorded steady-state responses remains difficult. The main problem is that the recorded response may not derive from a single generator. Many different regions of the auditory nervous system from the auditory nerve to the auditory cortex generate responses that follow the modulation signal of an amplitude-modulated tone. Each of these regions can create a response that has a specific phase and latency relationship to the stimulus. All these responses will overlap to make up the response recorded from the scalp. The phase and latency characteristics of the scalp-recorded response may be very complicated unless one of the generators is clearly dominant.

Changes in the phase of the 40 Hz changes in the level of arousal illustrate these phenomena. In a waking subject, the dominant response at 40 Hz probably derives from the auditory cortex, with the brainstem only contributing a little. With sleep, the cortical response may attenuate and the scalp-recorded response would then mainly reflect those potentials generated in the brainstem. The response during sleep is therefore much smaller than during wakefulness and occurs earlier (Lins and Picton, 1995).

In our particular case, the amplitude of the responses from different generators may vary differently with the carrier frequency. If so, the phase of the scalp-recorded response that combines potentials from the different brain generators, may not show a clear and regular relationship to carrier frequency. Furthermore, such data might provide an apparent latency that does not fit with an integer number of cycles prior to the recorded response. It is best only to calculate the apparent latency over regions where there is no change in the slope of the phase-frequency curve, since this suggests

one dominant generator. Different generators at different frequencies could cause the slope to change (e.g., for 750 Hz on the right of Fig. 4). Another reason for non-integer preceding cycles would be some undetermined filter delay in the system. These factors (multiple generators and unknown filters) may explain some of the discrepancies between our measurements of apparent latency and the estimated latencies obtained by postulating an integer number of cycles of modulation prior to the recorded response. Although the latencies were generally concordant, the apparent latencies measured for the 750 Hz carrier were longer than those estimated using preceding cycles (22.2 versus 20.7 ms at 80–100 Hz and 15.1 versus 11.3 ms at 150–190 Hz).

The advantages of using the 'preceding cycles' approach over apparent latency is that once the number of preceding cycles is estimated from using multiple stimulus rates, then only one rate is needed to make subsequent latency measurements. The disadvantage is that it has no clear way to handle the possibility of multiple generators, which could then lead to a non-integer number of preceding cycles. However, although the group delay or apparent latency measurements allow for such non-integer effects, the interpretation of the results in such a situation is not clear. In our results, the apparent latency measurements differed a little from the latencies obtained with the preceding cycles approach, particularly for the low-frequency stimuli (750 Hz). In this situation, one must consider the possibility of overlapping responses with different latencies, e.g., from different brainstem neurons.

We feel that the scalp-recorded potential in response to tones amplitude-modulated at frequencies 80–100 Hz probably derives mainly from a generator in the high brainstem. Preliminary source analyses in our laboratory and elsewhere (Mauer and Döring, 1999) support this idea. If this generator is clearly dominant over other generators, for example in the auditory nerve or lower brainstem, the derived latencies should be quite accurate. The relations of the response phase to carrier frequency should then reflect the effects of transport time and filter build-up in the cochlea. For the stimulation rates of 150–190 Hz, the responses that we are recording may reflect approximately equal contributions from generators in both the high brainstem and lower brainstem.

The measurement of latency using tones that are sinusoidally amplitude modulated is fraught with ambiguities. Another approach to estimating the latencies uses a modulating envelope that is pseudorandom rather than sinusoidal (Møller, 1985; Møller and Jho, 1989). Latencies of the response can then be calculated using a cross-correlation function between the response and the envelope of the stimulus. The pseudorandom envelope has to be within the frequency range that generates an envelope-following response. This has worked well for unit recordings in the brainstem (Møller, 1985) or for recordings from the human auditory nerve (Møller and Jho, 1989). The fact that we can record responses over a range of modulating frequencies makes this approach worth exploring for human recordings. However, the fact that many different regions of the brainstem may contribute to the scalp-recorded responses and that these regions may have different latencies and modulation tuning curves may make the responses difficult to interpret.

4.3. Latencies in the auditory system

Five main factors contribute to the latency of a neurophysiological response to an auditory stimulus:

(i) Acoustic delay. This is the delay between the stimulus being produced and the stimulus arriving at the oval window after transmission through the air to the tympanic membrane and through the middle ear.

(ii) Transport time. This is the delay between the arrival of the acoustic energy at the oval window and the beginning of activity at the location on the basilar membrane where transduction occurs for a stimulus of that particular frequency. High frequencies of sound maximally activate the basilar membrane close to the oval window and low frequencies maximally activate the regions of the basilar membrane at some distance from the oval window. The transport time therefore depends upon the frequency of the stimulus. Eggermont (1979b) estimated this 'mechanical' time (in ms) as $f^{-0.5}$ where f is the characteristic frequency (in kHz).

(iii) Filter built-up time. This is the time taken for the acoustic energy to pass through the active filtering process ('cochlear filter') of the hair cells that are sensitive to the frequencies of sound in the stimulus. This time will depend upon the sharpness of the filter, a parameter that can be measured by the Q_{10dB} parameter which is the characteristic frequency of the filter divided by the bandwidth of the filter at an intensity 10 dB above the threshold intensity for the filter. Values for Q_{10dB} in mammalian cochleae vary between 1 and 15 with the filters for lower frequency sounds having lower values (Eggermont, 1979b; Ruggero, 1992). The filter delay is usually measured in terms of the number of cycles to maximum amplitude and this varies with the Q_{10dB} with delays increasing by 2–3 ms as the stimulus frequency increases from 1 to 10 kHz (Eggermont, 1979b, data compensated by removing β as discussed later).

(iv) Synaptic delay. As the hair cells react to the sounds they synaptically activate the afferent nerve fibers. The transmission time across the synapse requires somewhere between 0.5 and 1.0 ms. This delay is not frequency-dependent.

(v) Conduction delay. This is the time taken for the

neural responses to travel from their point of origin on the afferent fiber terminals to the location from which they are recorded. For human evoked potentials this time includes some additional synaptic and conduction delays as the impulses travel to the place where they generate the response recorded from the scalp. The conduction time can be estimated from studies wherein the auditory nerve fibers are electrically stimulated. Shirane and Harrison (1991) found a delay of below 1 ms in the colliculus response as the characteristic frequency of the chinchilla afferent fiber being electrically stimulated decreased from 10 kHz to 1 kHz (compared to about 2 ms when using acoustic clicks).

The latency of the neural response is unfortunately not just a simple sum of these five delays. The problem centers on the temporal characteristics of neural responsiveness, particularly refractory times, and the way in which field potentials are generated. When presented with a transient stimulus the nerve fibers will respond more to the initial cycles of activation and less to later cycles. This relationship becomes even more complicated if one measures the combined responses of multiple neurons, such as the compound action potential of the auditory nerve. The neurons contributing to this combined response will show a significant variance in their response patterns and this can make the compound response quite different from the individual responses. One of the main effects is to accentuate the response at the onset of a stimulus. The discharges of many of the neurons will be relatively synchronous at the beginning of the stimulus and will become less synchronous as the stimulus continues (Elberling, 1976). Because of these factors, the latency of the neurophysiological response to a transient stimulus is not as long as the delay caused by the filter build-up. The response occurs some time during the build-up of the filter response and this time can vary with other parameters such as stimulus intensity. In order to account for this effect, Eggermont (1979b) multiplied the delay by a factor β between 0 and 1.0, in general estimating β at 0.5.

When recording steady-state responses, the neural response patterns have reached a balance between reactivity and refractoriness and this balance remains constant during the recording. We therefore suggest that for the steady-state response β is effectively equal to 1.0. This means that the delay to the peak of the response waveform varies in the same way as the delay to the peak of the changing stimulus (rather than a fraction of that stimulus delay). If so, the frequency-related delay of the steady-state response is exactly equal to the sum of the transport time and the filter build-up time for stimuli with frequencies equal to the carrier frequency.

Most studies of the neuronal responses to AM stimulation in animals (e.g., Rees and Palmer, 1989; Langner, 1992) have not considered the phase of the response patterns. Other studies (e.g., Møller, 1972; Rees and Møller, 1983) have measured changes in the phase of the discharge patterns with modulation frequency and how this relates to the temporal modulation transfer function of the neurons. Joris and Yin (1992) measured the phases of the neuronal responses to AM tones in the auditory nerve of the cat and found group delays related to carrier frequency increasing about 3 ms from 10 kHz to 500 Hz. This was similar to the changes in group delay estimated from the phase locked responses to pure tones (e.g., Goldstein et al., 1971; Gummer and Johnstone, 1984). Our data show a change of 5–6 ms over these frequencies, the larger delays probably being related to species differences in the basilar membrane and the travelling wave velocity, or some frequency-related change in conduction time (see v above).

4.4. Relations to otoacoustic emissions

Kimberley et al. (1993) studied frequency-related delays in cochlear processing using distortion product otoacoustic emissions (DPE). The latency of these responses is usually considered in terms of a 'round trip' delay: the acoustic signal must first reach the hair cells and then the acoustic energy generated by the hair cells must travel back to the external ear canal to be recorded up by the microphone. Dividing the DPE latency by 2 provides latencies that are roughly equivalent to the latencies of the neurophysiological responses. However, as Kimberley et al. (1993) clearly state, this is a simplified view. They point out that the latency of the DPE "consist[s] of (a) the actual travelling-wave delay to the f2 frequency place, (b) the cochlear filter delay at the site of generation, near f2, and (c) the backward travel time of the DPE from the generation site to the oval window." The reason for the rough equivalence between one half of the DPE latency and the frequency-related latency of the evoked potentials is that β is approximately 0.5 and the transport times (a and c) are small relative to the filter delay (b).

If the travelling wave delay is equal to the backward travel time, the latency of the DPE (t_e) can be represented by:

$$t_e = t_f + 2t_b$$

where t_f is the filter delay and t_b is the transport time on the basilar membrane. We are proposing that the frequency-related delay in the steady-state response (t_s) is represented by:

$$t_s = t_f + t_b$$

Combining these measurements could provide separate estimates for the transport time and the filter time. The latencies of the otoacoustic emissions change approximately 7 ms from 10 kHz to 1 kHz (Kimberley et al., 1993; Moulin and Kemp, 1996a). Our data indicate that the latencies of the steady-state responses would change by about 5 ms over the same frequency range. This would suggest that filter build-up causes frequency-related delays of up to 3 ms and that the transport time varies up to 2 ms.

Although illustrating how latency data for emissions and steady-state responses might be related, this analysis overly simplifies the nature of the DPE delay. First, the forward and backward transport times are not necessarily equal. Brown et al. (1994) review several different physiological models in the context of how DPE latencies change with development. The backward travel time may be very brief if it is based on a compression wave (cf., Sutton and Wilson, 1983). Moulin and Kemp (1996b) have presented evidence supporting the idea that the return time of the DPE is significantly shorter than its forward latency. A second major difficulty in the analysis concerns the location on the basilar membrane responsible for generating the DPEs. Moulin and Kemp (1996b) demonstrated that different DPEs (e.g., $2f_1-f_2$ or $2f_2-f_1$) are likely generated in different cochlear locations. Third, it is not clear that the return path is direct. Shera and Guinan, 1999 have recently reviewed our understanding of the delays measured for the DPEs. They suggest that the 'sources' of the backward-travelling waves are locations on the basilar membrane characterized by nonlinear electromechanical distortion. In this view, both stimulus frequency and distortion product emissions ultimately arise from linear reflection at these locations. The reflected energy derives from the stimulus-evoked travelling wave for stimulus-evoked emissions and from the nonlinear processing at regions of overlap between the f_1 and f_2 sounds for distortion product emissions.

Clearly, we shall have to understand the nature of the DPE latency more fully before we can combine it with the latency of the steady-state response to determine the relative contributions of transport and filtering to cochlear delays. Using the same stimuli to evoke both the emissions and the steady-state responses might decrease the number of variables and simplify the relations between these physiological responses. This approach would lead to examining the steady-state responses to cochlear distortion products (e.g., Chertoff and Hecox, 1990; Rickman and Chertoff, 1991) or the emissions evoked by amplitude-modulated tones.

4.5. Effects of intensity

The latency of wave V of the click-evoked auditory brainstem response increases from 5.5 ms to 8.5 ms as the intensity is decreased from 70 dB to 20 dB above threshold. Eggermont and Don (1980) considered this latency change in light of derived response studies. Part of the change can be attributed to a shift in the region of the cochlea that contributes most prominently to the response. At high intensities the basal end of the cochlea dominates but at low intensities the region of the cochlea mediating frequencies near 1000 Hz dominates. However, this can only account for about 1 ms of the change and there remains an intensity-related change of about 2 ms that occurs within each derived band response. Several factors contribute to this residual latency change. There might be a small place shift within the narrow band of frequencies responsible for the response. The β parameter also changes since clicks of higher intensity evoke responses earlier on the rise time of the synaptic activation pattern. There might be further effects of refractoriness and synchronization (cf. steps 3 and 4 in our model). Finally, there might be some increased latency in the brainstem pathways due to decreased efficiency of synaptic transmission when fewer afferent fibers are activated and firing of the post-synaptic cells is nearer threshold. The steady-state responses show a latency change of 2.4 ms from 75 to 35 dB SPL (roughly equivalent to 65 and 25 dB HL). We have suggested that steady-state responses are little affected by the β parameter and the effects of refractoriness and synchronization. This leaves some place shift in the cochlea or changes in synaptic transmission as possible causes of the intensity-related delay.

A place shift in the cochlea probably contributes significantly to the change in our latencies with intensity. Our stimuli were pure tones rather than broad-band stimuli and we did not try to eliminate the effects of place shift by using derived response masking techniques (which would be problematic with the multiple stimulus approach). Studies using the cochlear microphonic (Honrubia and Ward, 1968) and intracellular recordings from the hair cells (Zwislocki, 1991) have clearly shown that the activation pattern of a pure tone on the basilar membrane shifts towards the stapes with increasing intensity.

Latency changes with intensity can also be mediated by changes in rise time of the post-synaptic potential. With louder sounds, the synaptic potential rises more rapidly either because the release of synaptic transmitter at each synapse is greater or because of more synapses being simultaneously activated. The effects of intensity on the latency means that close attention must be paid to calibrating the stimuli. This is a particular problem when using some ear inserts since the intensities of the stimuli with carrier frequencies of 6000 Hz may be 20 dB lower than those of with lower carrier frequencies. In experiment 7, the data at 6000 Hz (Fig. 1) probably

over-estimate the actual latencies by 0.5–1.0 ms. One way to eliminate the confounds of intensity might be to use continuous tones which are partially amplitude modulated (Møller, 1975, 1985). The ongoing tone maintains the post-synaptic neuron near threshold and the modulation brings it immediately over or under the threshold. The latency then does not change significantly with the intensity of the tone since the post-synaptic neuron does not have to reach threshold from a resting level. Responses to transient stimuli show much larger latency effects because the silence between the stimuli allows the post-synaptic membrane to return to resting polarization levels. Our tones were 100% amplitude-modulated and may have been similar to transient stimuli in this regard. It will be interesting to see if we can decrease or remove the intensity effects using 30% amplitude-modulated tones.

The responses to carrier frequencies 500 and 750 Hz tended to be lower in amplitude than the others. We have obtained similar results in other experiments (John and Picton, 1999). Part of this effect might be caused by the lower effective intensity of these stimuli (since HL thresholds are a little higher at 500 and 750 Hz than at the higher frequencies). Another contributing factor may be the fact that these low carrier frequencies evoke an activation pattern on the basilar membrane that covers a greater spatial extent than higher carrier frequencies. Because neurons along a broad area of the basilar membrane all respond to the same low carrier frequency, some of the neurons responding might be activated significantly earlier than others. The resultant latency jitter of the responses might then attenuate the amplitude of the compound response.

4.6. Gender effects

The auditory brainstem response shows consistent latency differences with gender. Wave V elicited by clicks is in general 0.2 ms later in male subjects than in female subjects. When considering derived responses, male subjects show a longer latency difference between the responses of the basal turn and the responses from the apical turn. Don et al. (1993) interpreted these gender differences in terms of two effects. The first was related to faster cochlear delay times in females caused by the shorter length (and greater stiffness) of the basilar membrane (Sato et al., 1991). For the 700 Hz derived response this would account for 0.5 ms difference between male and female subjects. The second process accounting for a further 0.2 ms difference could be explained by gender differences in the size or the length of the auditory nerve, or by changes in the synchronization of the responses caused by the faster travelling wave velocity. Our findings for experiments 5 and 7 also suggested a gender-related difference in latency with male subjects having longer latencies than female subjects. However, the data contained considerable inter-subject variability and did not show any clear differences in this effect across the different carrier frequencies.

4.7. Effects of pathology

Measuring the phase and latency of auditory steady-state responses should increase our understanding of the underlying cochlear and neuronal mechanisms. It is also possible that such measurements may help evaluate patients with cochlear dysfunction. A main problem in such applications will be the reliability of the data. Close attention to relaxing the subjects and to eliminating artifacts improves the quality of the data. With experience, our intersubject variance has gone down (compare the standard deviations of Table 3 from earlier studies with those of Table 2). The latency estimates are quite stable from one measurement to the next (Table 4). However, the limits of normal (a normal subject may change up to 0.5 ms or more from one measurement to another) may be greater than the changes that occur due to changing clinical conditions (e.g., Munro et al. (1995) suggest that one might need to detect changes in the range of 0.3 ms when endolymphatic hydrops occurs or when hydrops is being treated).

In patients with endolymphatic hydrops there may be decreases in the transport time due to increased pressure levels across the basilar membrane. The resultant increase in stiffness of the basilar membrane will decrease the transport time. Thornton and Farrell (1991) used the derived auditory brainstem response to measure the latency (and hence the travelling wave velocity) at different locations on the basilar membrane. They found that the velocity was greater than normal in the high frequency regions of the cochlea in patients with endolymphatic hydrops. However, sensorineural hearing loss causes a broadening of the tuning curves of the auditory nerve fibers (e.g., Dallos and Harris, 1978) and this will decrease the filter-time and cause an earlier response (Eggermont, 1979a; Don et al., 1998). It is therefore important in Meniere's disease to disentangle the effects of transport time (due to endolymphatic hydrops) and filter time (hair cell damage). The latency changes in the physiological responses in patients with Meniere's disease seem to occur independently of filter changes that might be caused by hair cell damage (Thornton and Farrell, 1991; Kim et al., 1994; Donaldson and Ruth, 1996). Nevertheless, reliably measuring the small latency changes with the auditory brainstem response is technically very difficult (Munro et al., 1995) and electrocochleographic measurement of the relative size of the summating potential and the

action potential (e.g., Margolis et al., 1995) remains the most commonly used electrophysiological aid to the diagnosis of Meniere's disease. It is possible that measuring the latencies of the steady-state responses might provide some diagnostic help in patients with Meniere's disease. However, we have not yet determined whether the measurements will be sufficiently reliable for clinical use.

4.8. Summary

The latencies of the human auditory steady-state responses to AM tones change significantly and consistently with the carrier frequency, with the latencies being shorter for higher carrier frequencies. The latency changes 5.5–6.0 ms when the carrier frequency changes from 500 to 6000 Hz (e.g., Table 2). This change appears to be mainly due to two cochlear processes: the transport time for the acoustic energy to reach the responsive region of the basilar membrane and the filter buildup time of the hair cell transduction process.

Acknowledgements

This research was funded by the Medical Research Council of Canada. The authors would also like to thank James Knowles and the Baycrest Foundation for their support. Some of these data were presented at the meeting of the Neuroscience Society, Los Angeles, November, 1998, and at the meeting of the International Electric Response Audiometry Study Group in Tromsø, Norway, June, 1999. The authors appreciate the technical assistance of Patricia Van Roon, and the helpful discussion and advice provided by Martin Regan and Jos Eggermont.

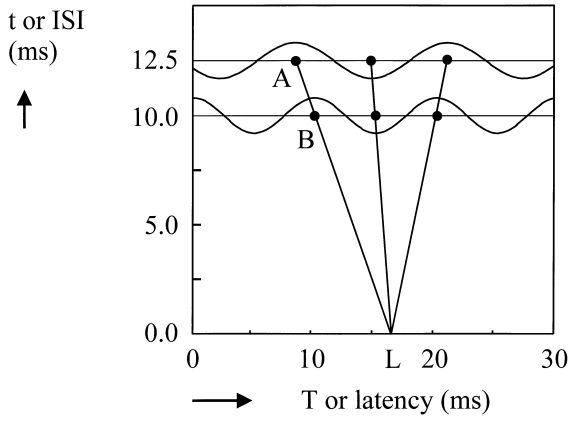

Fig. 14. Graphic estimation of apparent latency. This figure illustrates the technique of Diamond (1977) to estimate the apparent latency of a steady-state response. The responses are plotted on the y-axis according to the interstimulus interval. A line drawn between homologous peaks in these responses, e.g., A and B, is extended to the x-axis and the apparent latency (L) is estimated at the intercept.

Appendix. The equivalence of different methods for measuring latency

The basic technique of apparent latency is to measure the slope of the phase versus frequency plot (Regan, 1966, 1989, p. 42). The only constraints are that the slope should be consistent across the stimulus frequencies examined, and that effects of physiological filtering on the response phase should be compensated prior to the calculations (Regan, 1989, p. 29). Diamond (1977) proposed a simple graphic way to estimate apparent latency. This is illustrated schematically in Fig. 14 using modeled data. The latency of a particular peak in the steady-state response is plotted as T along the x-axis against the interstimulus interval plotted as t along the y-axis. A line joining the latencies is then extended to intercept the x-axis at a latency L equivalent to the apparent latency. Using two points (A and B) for simplicity, the slope of the line is

$$(t_B - t_A)/(T_B - T_A)$$

and the equation for the intercept is

$$L = T_B - t_B(T_B - T_A)/(t_B - t_A)$$

If we substitute into this equation

$$T_X = \theta_X/(360F_X)$$

and

$$t_X = 1/F_X$$

where θ is the phase delay of the response at the chosen peak and F is the frequency of stimulation, we can simplify the equation to obtain

$$L = (\theta_A - \theta_B)/(360(F_A - F_B))$$

which is the definition of apparent latency.

The technique described in this paper measures the latency as

$$L = \theta_A/(360F_A) + N/F_A$$

where N is an integer chosen so that both measurements (A and B) give the same value

$$\theta_A/(360F_A) + N/F_A = \theta_B/(360F_B) + N/F_B$$

This equation can be recast by bringing both phases over to the left, adding $\theta_B(F_A-F_B)/360$ to both sides and then dividing by $F_B(F_A-F_B)$ to give

$$(\theta_A-\theta_B)/(360(F_A-F_B)) = \theta_B/(360F_B) + N/F_B$$

This makes the measured latency (on the right) equal the apparent latency. The only difference between the two techniques is that apparent latency allows non-integer values of N.

References

Anderson, D.J., Rose, J.E., Hind, J.E., Brugge, J.F., 1971. Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects. J. Acoust. Soc. Am. 49, 1131–1139.

Azzena, G.B., Conti, G., Santarelli, R., Ottaviani, F., Paludetti, G., Maurizi, M., 1995. Generation of human auditory steady-state responses (SSRs). I: Stimulus rate effects. Hear. Res. 83, 1–8.

Basar, E., Rosen, B., Basar-Eroglu, C., Greitschus, F., 1987. The associations between 40 Hz-EEG and the middle latency response of the auditory evoked potential. Int. J. Neurosci. 33, 103–117.

Bijl, G.K., Veringa, F., 1985. Neural conduction time and steady-state evoked potentials. Electroencephalogr. Clin. Neurophysiol. 62, 465–467.

Bowman, D.M., Brown, D.K., Eggermont, J.J., Kimberley, B.P., 1997. The effect of sound intensity on f1-sweep and f2-sweep distortion product otoacoustic emissions phase delay estimates in human adults. J. Acoust. Soc. Am. 101, 1550–1559.

Brown, D., Kimberley, B., Eggermont, J., 1994. Cochlear traveling-wave delays estimated by distortion product emissions in normal hearing adults and term-born neonates. J. Otolaryngol. 23, 234–238.

Brown, P.B., Franz, G.N., Moraff, H., 1982. Electronics for the Modern Scientist. Elsevier, Amsterdam.

Carney, L.H., 1993. A model for the responses of low-frequency auditory-nerve fibers in cat. J. Acoust. Soc. Am. 93, 401–417.

Chertoff, M.E., Hecox, K.E., 1990. Auditory nonlinearites measured with auditory-evoked potentials. J. Acoust. Soc. Am. 87, 1248–1254.

Cohen, L.T., Rickards, F.W., Clark, G.M., 1991. A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. J. Acoust. Soc. Am. 90, 2467–2479.

Corey, D.P., Hudspeth, A.J., 1983. Analysis of the microphonic potential of the bullfrog's sacculus. J. Neurosci. 3, 942–961.

Dallos, P., 1985. Response characteristics of mammalian cochlear hair cells. J. Neurosci. 5, 1591–1608.

Dallos, P., Cheatham, M.A., 1971. Travel time in the cochlea and its determination from cochlear-microphonic data. J. Acoust. Soc. Am. 49, 1140–1143.

Dallos, P., Harris, D., 1978. Properties of auditory nerve responses in absence of outer hair cells. J. Neurophysiol. 41, 365–383.

Dancer, A., 1992. Experimental look at cochlear mechanics. Audiology 31, 301–312.

Diamond, A.L., 1977. Latency of the steady state visual evoked potential. Electroencephalogr. Clin. Neurophysiol. 42, 125–127.

Dobie, R.A., Wilson, M.J., 1993. Objective response detection in the frequency domain. Electroencephalogr. Clin. Neurophysiol. 88, 516–524.

Dobie, R.A., Wilson, M.J., 1996. A comparison of t test, F test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids. J. Acoust. Soc. Am. 100, 2236–2246.

Dobie, R.A., Wilson, M.J., 1998. Low-level steady-state auditory evoked potentials: effects of rate and sedation on detectability. J. Acoust. Soc. Am. 104, 3482–3488.

Don, M., 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., Ponton, C.W., Eggermont, J.J., 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.

Don, M., Ponton, C.W., Eggermont, J.J., Kwong, B., 1998. The effects of sensory hearing loss on cochlear filter times estimated from ABR latencies. J. Acoust. Soc. Am. 104, 2280–2289.

Donaldson, G.S., Ruth, R.A., 1996. Derived-band auditory brainstem response estimates of traveling wave velocity in humans: II. Subjects with noise-induced hearing loss and Meniere's disease. J. Speech Hear. Res. 39, 534–545.

Eggermont, J.J., 1979a. Compound action potentials: tuning curves and delay times. Scand. Audiol. Suppl. 9, 129–139.

Eggermont, J.J., 1979b. Narrow-band AP latencies in normal and recruiting human ears. J. Acoust. Soc. Am. 65, 463–470.

Eggermont, J.J., 1985. Peripheral auditory adaptation and fatigue: A model oriented review. Hear. Res. 18, 57–71.

Eggermont, J.J., 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.

Eggermont, J.J., Brown, D.K., Ponton, C.W., Kimberley, B.P., 1996. Comparison of distortion product otoacoustic emission (DPOAE) and auditory brain stem response (ABR) traveling wave delay measurements suggests frequency-specific synapse maturation. Ear Hear. 17, 386–394.

Elberling, C., 1976. Modeling action potentials. Rev. Laryngol. Otol. Rhinol. (Bord.) Suppl. 97, 527–537.

Fisher, N.I., 1993. Statistical Analysis of Circular Data. Cambridge University Press, Cambridge.

Fridman, J., Zappulla, R., Bergelson, M., Greenblatt, E., Malis, L., Morrell, F., Hoeppner, T., 1984. Application of phase spectral analysis for brain stem auditory evoked potential detection in normal subjects and patients with posterior fossa tumors. Audiology 23, 99–113.

Galambos, R., Makeig, S., Talmachoff, P.J., 1981. A 40-Hz auditory potential recorded from the human scalp. Proc. Natl. Acad. Sci. USA 78, 2643–2647.

Glasberg, B.R., Moore, B.C., 1990. Derivation of auditory filter shapes from notched-noise data. Hear. Res. 47, 103–138.

Goldstein, J.L., Baer, T., Kiang, N.Y.S., 1971. A theoretical treatment of latency, group delay, and tuning characteristics for auditory-nerve responses to clicks and tones. In: Sachs, M.B. (Ed.), Physiology of the Auditory System; Based on the Proceedings of a Workshop. National Educational Consultants, Baltimore, MD, pp. 133–141.

Gummer, A.W., Johnstone, B.M., 1984. Group delay measurement from spiral ganglion cells in the basal turn of the guinea pig cochlea. J. Acoust. Soc. Am. 76, 1388–1400.

Gutschalk, A., Mase, R., Roth, R., Ille, N., Rupp, A., Hähnel, S., Picton, T.W., Scherg, M., 1999. Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clin. Neurophysiol. 110, 856–868.

Hari, R., Hamalainen, M., Joutsiniemi, S.L., 1989. Neuromagnetic steady-state responses to auditory stimuli. J. Acoust. Soc. Am. 86, 1033–1039.

Honrubia, V., Ward, P.H., 1968. Longitudinal distribution of the

cochlear microphonics inside the cochlear duct (guinea pig). J. Acoust. Soc. Am. 44, 951–958.
John, M.S., Picton, T.W., 2000. MASTER: A Windows program for recording multiple auditory steady-state responses. Comput. Methods Programs Biomed. 61, 125–150.
John, M.S., Lins, O.G., Boucher, B.L., Picton, T.W., 1998. Multiple auditory steady-state responses (MASTER): stimulus and recording parameters. Audiology 37, 59–82.
Joris, P.X., Yin, T.C., 1992. Responses to amplitude-modulated tones in the auditory nerve of the cat. J. Acoust. Soc. Am. 91, 215–232.
Kim, Y., Aoyagi, M., Koike, Y., 1994. Measurement of cochlear basilar membrane traveling wave velocity by derived ABR. Acta Otolaryngol. Suppl. 511, 71–76.
Kimberley, B.P., Brown, D.K., Eggermont, J.J., 1993. Measuring human cochlear traveling wave delay using distortion product emission phase responses. J. Acoust. Soc. Am. 94, 1343–1350.
Langner, G., 1992. Periodicity coding in the auditory system. Hear. Res. 60, 115–142.
Levi, E.C., Folsom, R.C., Dobie, R.A., 1993. Amplitude-modulation following response (AMFR): effects of modulation rate, carrier frequency, age, and state. Hear. Res. 68, 42–52.
Lins, O.G., Picton, T.W., 1995. Auditory steady-state responses to multiple simultaneous stimuli. Electroencephalogr. Clin. Neurophysiol. 96, 420–432.
Lins, O.G., Picton, P.E., Picton, T.W., Champagne, S.C., Durieux-Smith, A., 1995. Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz. J. Acoust. Soc. Am. 97, 3051–3063.
Lins, O.G., Picton, T.W., Boucher, B.L., Durieux-Smith, A., Champagne, S.C., Moran, L.M., Perez-Abalo, M.C., Martin, V., Savio, G., 1996. Frequency-specific audiometry using steady-state responses. Ear Hear. 17, 81–96.
Margolis, R.H., Rieks, D., Fournier, E.M., Levine, S.E., 1995. Tympanic electrocochleography for diagnosis of Meniere's disease. Arch. Otolaryngol. Head Neck Surg. 121, 44–55.
Mauer, G., Döring, W.H., 1999. Generators of amplitude modulation following response (AMFR). Paper presented at the June 1999 meeting of the International Evoked Response Audiometry Study Group, Tromsø.
Møller, A.R., 1972. Coding of amplitude and frequency modulated sounds in the cochlear nucleus of the rat. Acta Physiol. Scand. 86, 223–238.
Møller, A.R., 1975. Latency of unit responses in cochlear nucleus determined in two different ways. J. Neurophysiol. 38, 812–821.
Møller, A.R., 1985. Origin of latency shift of cochlear nerve potentials with sound intensity. Hear. Res. 17, 177–189.
Møller, A.R., Jho, H.D., 1989. Responses from the exposed human auditory nerve to pseudorandom noise. Hear. Res. 42, 237–252.
Moulin, A., Kemp, D.T., 1996a. Multicomponent acoustic distortion product otoacoustic emission phase in humans. I. General characteristics. J. Acoust. Soc. Am. 100, 1617–1639.
Moulin, A., Kemp, D.T., 1996b. Multicomponent acoustic distortion product otoacoustic emission phase in humans. II. Implications for distortion product otoacoustic emissions generation. J. Acoust. Soc. Am. 100, 1640–1662.
Munro, K.J., Smith, R., Thornton, A.R., 1995. Difficulties experienced in implementing the ABR travelling wave velocity (Delta V) technique with two commercially available systems. Br. J. Audiol. 29, 23–29.
Neely, S.T., Norton, S.J., Gorga, M.P., Jesteadt, W., 1988. Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli. J. Acoust. Soc. Am. 83, 652–656.
O'Mahoney, C., Kemp, D.T., 1995. Distortion product otoacoustic emission delay measurement in human ears. J. Acoust. Soc. Am. 97, 1–15.
Parker, D.J., Thornton, A.R., 1978. Cochlear travelling wave velocities calculated from the derived components of the cochlear nerve and brainstem evoked responses of the human auditory system. Scand. Audiol. 7, 67–70.
Patterson, R., 1994. The sound of a sinusoid: spectral models. J. Acoust. Soc. Am. 96, 1409–1418.
Patuzzi, R., 1996. Cochlear micromechanics and macromechanics. In: Dallos, P., Popper, A.N., Fay, R.R. (Eds.), Springer Handbook of Auditory Research, Vol. 8. The Cochlea. Springer, New York, pp. 187–223.
Picton, T.W., Ouellette, J., Hamel, G., Smith, A.D., 1979. Brainstem evoked potentials to tonepips in notched noise. J. Otolaryngol. 8, 289–314.
Plourde, G., Stapells, D.R., Picton, T.W., 1991. The human auditory steady-state evoked potentials. Acta Otolaryngol. Suppl. 491, 153–159.
Rance, G., Rickards, F.W., Cohen, L.T., De Vidi, S., Clark, G.M., 1995. The automated prediction of hearing thresholds in sleeping subjects using auditory steady-state evoked potentials. Ear Hear. 16, 499–507.
Rees, A., Møller, A.R., 1983. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. Hear. Res. 10, 301–330.
Rees, A., Palmer, A.R., 1989. Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. J. Acoust. Soc. Am. 85, 1978–1994.
Regan, D., 1966. Some characteristics of average steady-state and transient responses evoked by modulated light. Electroencephalogr. Clin. Neurophysiol. 20, 238–248.
Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier, New York.
Rickards, F.W., Clark, G.M., 1984. Steady state evoked potentials to amplitude-modulated tones. In: Nodar, R.H., Barber, C. (Eds.), Evoked Potentials II. Butterworth, Boston, MA, pp. 163–168.
Rickards, F.W., Tan, L.E., Cohen, L.T., Wilson, O.J., Drew, J.H., Clark, G.M., 1994. Auditory steady-state evoked potential in newborns. Br. J. Audiol. 28, 327–337.
Rickman, M.D., Chertoff, M.E., 1991. Electrophysiological evidence of nonlinear distortion products to two-tone stimuli. J. Acoust. Soc. Am. 89, 2818–2826.
Rodriguez, R., Picton, T., Linden, D., Hamel, G., Laframboise, G., 1986. Human auditory steady state responses: effects of intensity and frequency. Ear Hear. 7, 300–313.
Ruggero, M.A., 1992. Physiology and coding of sound in the auditory nerve. In: Popper, A.N., Fay, R.R. (Eds.), Springer Handbook of Auditory Research, Vol. 2. The Mammalian Auditory Pathway: Neurophysiology. Springer-Verlag, New York, pp. 35–93.
Ruggero, M.A., 1994. Cochlear delays and traveling waves: comments on 'Experimental look at cochlear mechanics'. Audiology 33, 131–142.
Santarelli, R., Maurizi, M., Conti, G., Ottaviani, F., Paludetti, G., Pettorossi, V.E., 1995. Generation of human auditory steady-state responses (SSRs). II: Addition of responses to individual stimuli. Hear. Res. 83, 9–18.
Sato, H., Sando, I., Takahashi, H., 1991. Sexual dimorphism and development of the human cochlea. Computer 3-D measurement. Acta Otolaryngol. 111, 1037–1040.
Schneider, S., Prijs, V.F., Shoonhovan, R., 1999. Group delays of distortion product otoacoustic emissions in the guinea pig. J. Acoust. Soc. Am. 105, 2722–2730.
Shera, C.A., Guinan, J.J., Jr., 1999. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. J. Acoust. Soc. Am. 105, 782–798.
Shirane, M., Harrison, R.V., 1991. The effects of long and short term profound deafness on the responses of inferior colliculus to elec-

trical stimulation of the cochlea. Acta Otolaryngol. Suppl. 489, 32–40.

Stapells, D.R., Linden, D., Suffield, J.B., Hamel, G., Picton, T.W., 1984. Human auditory steady state potentials. Ear Hear. 5, 105–113.

Stapells, D.R., Makeig, S., Galambos, R., 1987. Auditory steady-state responses: threshold prediction using phase coherence. Electroencephalogr. Clin. Neurophysiol. 67, 260–270.

Sutton, G.J., Wilson, J.P., 1983. Experiments in hearing. Modelling cochlear echoes: the influence of irregularities in frequency mapping on summed cochlear activity. In: de Boer, E., Viergever, M. (Eds.), Mechanics of Hearing. Nijhoff, The Hague, pp. 83–90.

Teas, D.C., Eldredge, D.H., Davis, H., 1962. Cochlear responses to acoustic transients: an interpretation of whole-nerve action potentials. J. Acoust. Soc. Am. 34, 1438–1459.

Thornton, A.R., Farrell, G., 1991. Apparent travelling wave velocity changes in cases of endolymphatic hydrops. Scand. Audiol. 20, 13–18.

Valdes, J.L., Perez-Abalo, M.C., Martin, V., Savio, G., Sierra, C., Rodriguez, E., Lins, O., 1997. Comparison of statistical indicators for the automatic detection of 80 Hz auditory steady state responses. Ear Hear. 18, 420–429.

van der Tweel, L.H., Verduyn Lunel, H.F.E., 1965. Human visual responses to sinusoidally modulated light. Electroencephalogr. Clin. Neurophysiol. 18, 587–598.

Von Békésy, G., 1960. Experiments in Hearing. McGraw-Hill, New York.

Von Békésy, G., 1963. Hearing theories and complex sounds. J. Acoust. Soc. Am. 35, 588–601.

Zurek, P.M., 1992. Detectability of transient and sinusoidal otoacoustic emissions. Ear Hear. 13, 307–310.

Zwislocki, J.J., 1991. What is the cochlear place code for pitch? Acta Otolaryngol (Stockh). 111, 256–262.