# Original Article

*Terence W. Picton*
*M. Sasha John*
*Andrew Dimitrijevic*
*David Purcell*

Rotman Research Institute,
Baycrest Centre for Geriatric Care,
University of Toronto, Canada

# Human auditory steady-state responses

## Respuestas auditivas de estado estable en humanos

## Abstract

Steady-state evoked potentials can be recorded from the human scalp in response to auditory stimuli presented at rates between 1 and 200 Hz or by periodic modulations of the amplitude and/or frequency of a continuous tone. Responses can be objectively detected using frequency-based analyses. In waking subjects, the responses are particularly prominent at rates near 40 Hz. Responses evoked by more rapidly presented stimuli are less affected by changes in arousal and can be evoked by multiple simultaneous stimuli without significant loss of amplitude. Response amplitude increases as the depth of modulation or the intensity increases. The phase delay of the response increases as the intensity or the carrier frequency decreases. Auditory steady-state responses are generated throughout the auditory nervous system, with cortical regions contributing more than brainstem generators to responses at lower modulation frequencies. These responses are useful for objectively evaluating auditory thresholds, assessing suprathreshold hearing, and monitoring the state of arousal during anesthesia.

## Sumario

Los potenciales evocados de estado estable pueden registrarse del cráneo humano en respuesta a estímulos auditivos presentados a tasas de 1 y 200 Hz o por modulaciones periódicas de la amplitud y/o de la frecuencia de un tono continuo. Las respuestas pueden ser detectadas objetivamente por medio de un análisis frecuencial En sujetos en estado de alerta las respuestas son particularmente prominentes con tasas de estimulación cercanas a 40 Hz. Las respuestas evocadas por estímulos presentados a tasa más rápida resultan menos afectadas por cambios del estado de conciencia y pueden ser evocados por estímulos múltiples simultáneos sin una pérdida significativa de la amplitud. La amplitud de la respuesta aumenta conforme la profundidad de la modulación o de la intensidad aumenta. El retraso de fase de la respuesta aumenta conforme la intensidad de la frecuencia portadora aumenta. Las respuestas auditivas de estado estable se generan a todo lo largo del sistema nervioso auditivo; las regiones corticales contribuyen más que los generadores del tallo cerebral en las respuestas de frecuencias más bajas. Estas respuestas son útiles para evaluar objetivamente los umbrales de audición y permiten también evaluar la audición supraliminar y monitorizar el estado de conciencia durante la anestesia.

## Introduction

This paper reviews the current state of our knowledge concerning the human auditory steady-state responses. After a brief introduction, the paper considers how these responses can be distinguished from the other activity recorded from the scalp. The paper then reviews how different auditory stimuli evoke the responses, and which regions of the brain respond to these stimuli. The effects of different subject parameters are then considered as a prelude to reviewing the clinical applications of the responses in audiometry, neurology, and anesthesia. The intent of the review is to demonstrate both how much we already know about these responses and how much we still need to find out.

A steady-state response is an evoked potential 'whose constituent discrete frequency components remain constant in amplitude and phase over an infinitely long time period' (Regan, 1989). Since infinity lies beyond our measurement, the response is considered stable over a temporal window much longer than the duration of a single stimulus cycle. This qualification of the definition allows us to measure slow changes (or unsteadiness) in the steady-state response. Steady-state responses are recorded when stimuli are presented periodically, and demonstrate how the brain 'follows' a stimulus or how the stimulus 'drives' a response.

Steady-state responses are classically distinguished from transient responses, the frequency components of which change

over time. In general, transient responses are evoked by stimuli that occur infrequently (in theory, only once), and steady-state responses are evoked by stimuli that occur more rapidly. Physiologic responses are generally best evaluated as steady-state responses when the stimulus rate is sufficiently rapid that the transient response to one stimulus overlaps with the responses to succeeding stimuli. If the brain responded in a linear fashion, steady-state responses would be completely predictable from the transient response. However, the brain is not linear, and steady-state and transient responses therefore provide independent views of its function.

Human steady-state evoked potentials were originally recorded in response to visual stimuli (Regan, 1966). These responses can often be seen in the ongoing electroencephalogram (EEG) as the photic driving response. Acoustic driving responses are not recognizable in the ongoing EEG, although they can be recorded from depth electrodes (Chatrian et al, 1960). Measuring these responses from the scalp required the development of averaging procedures to attenuate the background EEG (Geisler, 1960). Geisler (1960) recorded responses to clicks presented at rates from 1 to 120 per second. At the faster rates, the responses contained most of their energy at frequencies equal to the rate of stimulation. In reviewing techniques for evoked response audiometry, Schimmel et al (1974) mentioned that the auditory responses with peak latencies between 20 and 40 ms could be efficiently recorded at stimulus rates of 40–45 Hz. They also suggested that the responses might then be identified after Fourier analysis by comparing the amplitude of the response at the frequency of stimulation to the amplitude at adjacent frequencies. Campbell et al (1977) recorded responses evoked by clicks, sinusoidally modulated tones and square-wave-modulated tones using stimulus rates between 8 and 32 Hz. The responses were larger at 12–16 Hz than at higher rates, and were larger for clicks and square-wave-modulated 500-Hz tones than for sinusoidally modulated 500-Hz or 2000-Hz tones. Hall (1979) recorded responses to amplitude-modulated tones composed by adding together tones of different frequencies and amplitudes, and demonstrated that the response followed the envelope of the sound rather than the frequency of the tone or the perceived periodicity pitch.

Galambos et al (1981) provided the main trigger for the extensive research into human auditory steady-state responses by showing that these were very prominent at stimulus rates near 40 Hz. They suggested that the 40-Hz potential is largely the result of the superposition of transient middle-latency responses. Since the peaks of the transient responses are separated in time by about 25 ms, this superposition would be most effective at 40 Hz. They used brief tone-bursts and demonstrated that the response was larger for lower than for higher tonal frequencies, and that the response could be recorded at intensities near hearing threshold. They also found that the response was much smaller when the subject was drowsy or asleep (Galambos, 1981).

Rickards and Clark (1984) showed that auditory steady-state responses can be recorded at many different stimulus rates, and that the amplitude of the responses decreased with increasing rate. Cohen et al (1991) found that sleep did not attenuate the response at modulation rates greater than 70 Hz. Furthermore, since the background noise decreased with increasing frequency, the signal-to-noise ratio at these high frequencies was higher than or equivalent to that found using slower modulation rates.

## Measurement

### Time and frequency domains

Any waveform that varies over time can be transformed into a spectrum of activity over different frequencies, and vice versa. In the time domain, a recording can be measured by selecting peaks and troughs and calculating their amplitudes and latencies. In the frequency domain, activity is measured as the amplitudes and phases at particular frequencies. Both transient and steady-state responses can be measured in either the time or frequency domain. Although the frequency spectrum of a transient response may be evaluated over the whole duration of the response, the transient nature of the response is best demonstrated by showing how the spectrum changes over time in a spectrogram. The response is more appropriately measured in such a spectrogram or in the time domain (as peaks and troughs).

Averaging multiple responses together in order to decrease the background noise of a recording blurs the distinction between transient and steady-state responses. When the stimulus rate is slow and irregular, the average response can be considered a reasonable representation of the transient response, although changes across the averaging period will be lost. A response evoked by a stimulus presented periodically at a slow rate will contain harmonics at the rate of stimulation. After some initial settling, the response can reach steady state. The response is still perhaps most easily evaluated in the time domain, but the measurement of the response can consider only activity with frequencies at the harmonics of the stimulus rate, and can discard the activity at intervening frequencies.

When a stimulus is presented at a rapid periodic rate, the spectrum of the response will show activity only at the stimulus rate and its harmonics. Since the high frequencies in a response are generally limited by low-pass filtering in either the brain or the recording device, the number of components in the spectrum is lower when the rate is higher. The response is most appropriately measured in the frequency domain, since one need only consider those peaks in the spectrum at the stimulus rate or its harmonics. However, in the time domain, the small number of spectral components translates into a simple periodic waveform, which can be measured in terms of peaks and troughs. This is not as accurate as in the frequency domain, since noise can distort the peak measurements, and since small changes in phase between harmonics can dramatically alter the peaks in the time waveform.

### Frequency-domain measurements

Steady-state responses are therefore generally measured in the frequency domain. Human auditory steady-state responses have been assessed using a Fourier analyzer (Stapells et al, 1984) or a Fourier transform (Rickards and Clark, 1984).

The Fourier analyzer (Regan, 1966, 1977, 1989) measures the response at a particular frequency by separately multiplying the recorded activity by the sine and the cosine of the stimulus frequency. The resultant products are low-pass filtered (or integrated over time) to give two sustained values, $x$ from the multiplication with the cosine, and $y$ from the multiplication

with the sine. From these values, the amplitude ($a$) and phase ($\theta$) of the response can be calculated:

$$a = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1}(y/x)$$

These calculations can be implemented in either analog or digital form. One advantage of the Fourier analyzer over the now more common Fourier transform is that variations in the frequency of the signal need not significantly affect the measurements, provided that the sine and cosine signals follow these variations.

The Fourier transform converts a set of values that change in amplitude over time to a set of complex values that vary with frequency. The complex values can then be converted to amplitudes and phases using the same formula as described in the previous paragraph, with $x$ being real and $y$ imaginary. Basically, the calculations described in the previous paragraph are performed for a set of frequencies, only one of which is the stimulus frequency. The transformation is usually implemented in digital form, most commonly using a fast Fourier transform (FFT) algorithm (Cooley and Tukey, 1965). This calculates specific real and imaginary values for a set of frequency bins ranging between 0 and half the analog–digital conversion rate (the Nyquist frequency), with each bin separated from the adjacent bin by a resolution equal to the reciprocal of the duration of the recording sweep. The FFT analysis of a response to a tone amplitude modulated in the 40-Hz range is shown in Figure 1. The FFT requires that the stimuli be exactly synchronized with the analog–digital conversion, and that the number of values in the time domain be a power of 2. The advantage of the FFT over the Fourier analyzer is that it provides simultaneous estimates of the activity at a whole spectrum of frequencies, in addition to the activity at one particular stimulus frequency.
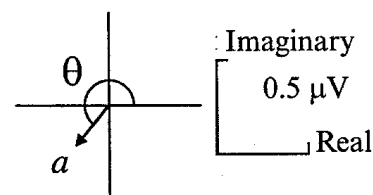
### Amplitude and phase

The literature presents many different measurements of the amplitude of steady-state responses. The early studies using visual analysis generally reported peak-to-peak amplitudes.
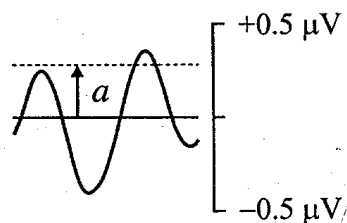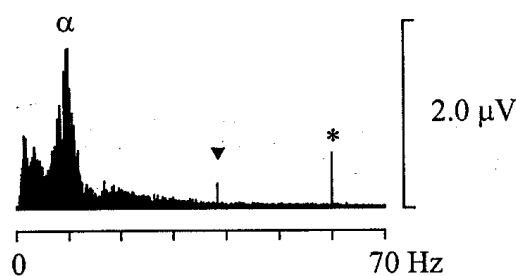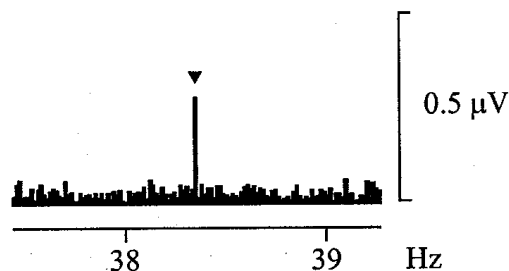


**Figure 1.** Recording and analysis of auditory steady-state responses. A continuous AM tone was presented at 70 dB SPL to the left ear through an insert earphone. The carrier frequency was 1000 Hz, the modulation frequency was 38.33 Hz (157 cycles in 4.096 s), and the depth of modulation was 100%. The upper left shows the response from one subject, recorded in the time domain from the vertex relative to an average reference (based on 47 scalp electrodes). Averaging was carried out using a sweep that was made equivalent to two cycles in duration by using spline-fitting. Below this is a polar plot of the response (the real part of the response is plotted on the $x$-axis and the imaginary part on the $y$-axis). The amplitude of the response ($a$) is shown by the length of the vector arrow, and the phase ($\theta$) by the angle that this vector makes with the $x$-axis. On the upper right is shown the amplitude spectrum from an FFT of a recording that lasted for 65.536 s. This shows EEG activity at low frequencies with a prominent peak at 10 Hz— the alpha ($\alpha$) rhythm—and some activity at the powerline frequency of 60 Hz (*). There is also a specific peak (indicated by the triangle) at the modulation frequency of the stimulus—the steady-state response. Below this is shown a close-up of a small region of the spectrum near the response frequency, with the response standing out above the adjacent frequencies. Comparing the response amplitude to the amplitude at these adjacent frequencies is the basis of the F-test.

Different FFT algorithms provide a variety of amplitude measurements, such as the peak-to-peak amplitude or the root-mean-square (rms) amplitude. In this paper, we shall measure the amplitude of the response as the baseline-to-peak amplitude, since this is what is normally used in the equations for a sine wave. When quoting the literature, we shall convert data to baseline-to-peak measurements. The size of the response can also be measured in terms of power (the square of the amplitude), particularly if one is considering the signal-to-noise ratio of the recording.

Measurements of phase also vary. Most analysis routines give the phase of the activity at the beginning of the recording (the 'onset' phase) and consider the phase relative to a cosine (the cosine being the measurement along the real axis for the complex numbers of the FFT). In physiologic applications, one is often more interested in the phase 'lag' or 'delay', which is calculated by subtracting the onset phase of the response from the onset phase of the stimulus. If the onset phase of the stimulus is zero, the phase delay is obtained by subtracting the onset phase of the response from 360°. The phase of a response depends on the polarity convention of the recording—if the response is inverted, the phase changes by 180°.

The circularity of the phase measurement must be taken into account when combining phases or estimating their variance (Fisher, 1993; Zar, 1999). 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°.

Steady-state responses may be plotted in several ways. A polar plot shows the response as a vector, extending outwards from the origin where the axes intersect. The amplitude of the response is the length of the line, and the phase is the angle between the line and the $x$-axis measured counter-clockwise from the axis. If one plots onset phase on a polar plot, and if the stimulus phase is zero, the phase delay is the angle measured clockwise (provided that both the stimulus and the response are measured in the same way, e.g. relative to cosine phase). The results of the FFT can also be plotted as a spectrum of amplitudes or powers, with the response at the stimulus frequency standing out above the other measurements. Figure 1 illustrates both polar plots and spectra.

*Phase and latency*
Phase delay (in degrees) can be converted to latency (in seconds) by dividing by $360f$, where $f$ is the stimulus modulation frequency (in Hz). Unfortunately, any direct conversion of phase to latency is fraught with problems (John and Picton, 2000a). First, filtering processes in both the presentation of the sound and the recording of the response can significantly affect the measured phase. Second, filtering processes occurring in the activation of the nervous system and the generation of the response can also affect the phase.

A major ambiguity in assessing phase is related to the steady-state nature of the response. The circularity of phase makes it difficult to determine whether the actual latency from the stimulus includes additional full cycles as well as the measured phase delay. If the phase lag of a particular response is 350° for a subject in one recording and 10° in a separate recording, it is most likely that the second response occurs after an additional
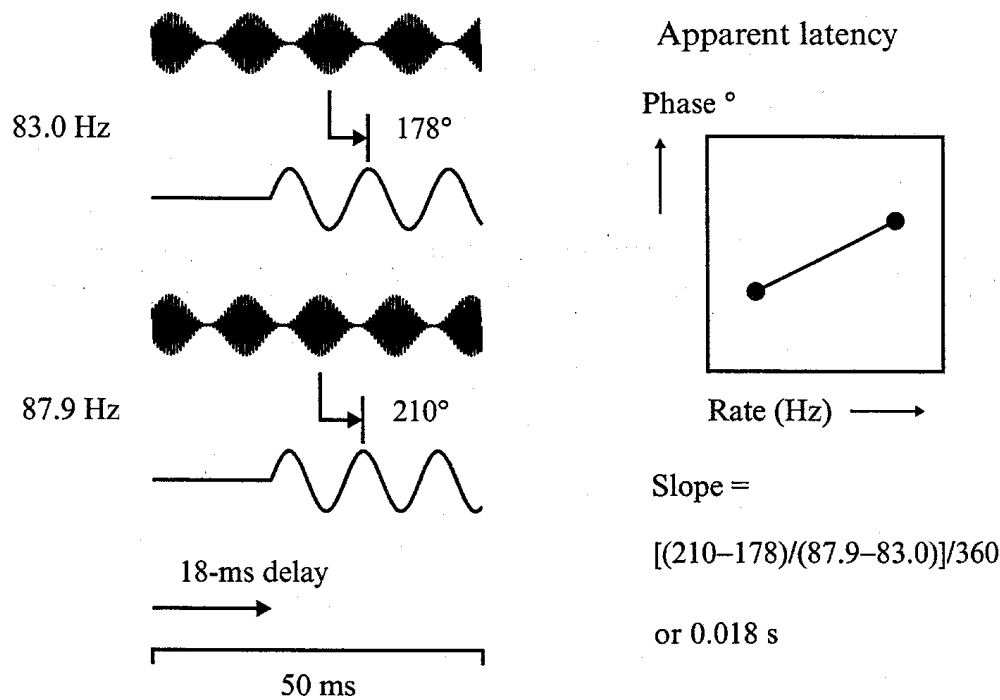
preceding cycle, i.e. at 370°. It is impossible, from one measurement, to determine how many cycles occurred between the presentation of the stimulus and the time at which the response was recorded.

Measuring the phase of the response at different stimulus frequencies can provide an estimate of latency that takes into account the possibility of preceding cycles and other delays. Regan (1966) plotted phase delay ($\theta$) against the stimulus frequency ($f$) and defined the slope, $\delta\theta/(360\delta f)$, as the 'apparent latency'. Recording multiple responses simultaneously (see 'Multiple simultaneous stimuli' below) can rapidly provide the multiple phase measurements for calculating apparent latency (Regan, 1976). Diamond (1977) proposed a way to make a similar measurement from data in the time domain rather than the frequency domain. In the engineering literature, apparent latency has been called 'group delay' (Goldstein et al, 1971).

In order for these measurements to be meaningful, the slope must be constant across the stimulus frequencies that are being examined. If there are changes in the slope, the responses may have been produced by more than one generator, each with its own latency. Furthermore, the generation of the response should not involve significant filtering across the stimulus frequencies, since filtering will change the phase of a response. This might be ensured by demonstrating a stable amplitude across the stimulus frequencies. If the amplitude does change, one can estimate the filtering process and compensate for its effects on phase (Regan, 1989). However, significant filtering during the generation of the response can render the measurement of apparent latency meaningless, unless the characteristics of the filtering are precisely known (Bijl and Veringa, 1985).

John and Picton (2000a) proposed a 'preceding-cycles' technique to measure the phase of a response when it has been recorded at multiple rates. This technique assumes that there is no delay in the response caused by filtering within the modulation frequencies, and that there is therefore an integer number of preceding cycles before the recorded response. It is essentially the same technique as apparent latency, except for the proviso that the response delay has to be an integer number of cycles.

Figure 2 shows diagrammatically how apparent latency is calculated. There has been some controversy in the literature concerning the meaning of apparent latency (Hari et al, 1989). As well as the previously mentioned filtering effects, there is a problem if the response is composed of more than one time-locked frequency component. If the measured scalp potential is dominated by a single source, then apparent latency will give the delay of the signal from that source with respect to the stimulus. The situation is less straightforward to interpret when there are multiple sources summing to form the net measured response. For example, the 40-Hz response may be derived from sources in both the brainstem and the cortex. If two sources sum relatively in phase (i.e. do not cancel one another out), then the calculated apparent latency will fall between the latencies of the two sources. The latency estimate will be weighted towards the largest source as a function of its relative strength. If the two sources create fields that tend to sum destructively at the electrode (near 180° out of phase, and of similar magnitude), the response signal will be very small and the apparent latency will be strongly weighted towards the largest source, even if its relative strength is only slighter greater than that of the next

**83.0 Hz**    178°

**87.9 Hz**    210°

18-ms delay

50 ms

Apparent latency

Phase °

Rate (Hz) ⟶

Slope =

[(210–178)/(87.9–83.0)]/360

or 0.018 s

**Figure 2.** Apparent latency. This figure illustrates diagrammatically the basic concept behind apparent latency. On the left are shown two AM stimuli with modulation frequencies of 83.0 and 87.9 Hz. These stimuli evoke a response that occurs with a delay of 18 ms. The responses are evaluated some time after the responses have begun (and after they have settled down into a steady-state—in this diagram, they immediately attain steady state). A phase lag between the stimulus and the response is then measured for each of the stimuli. These phase lags can be plotted against the modulation frequency and the delay estimated from the slope of the graph.

largest source. Apparent latency is a useful tool for indirectly estimating delays in the auditory system, but it must be used and interpreted with caution.

## Signals in noise

*Recognizing signals*
Human auditory steady-state responses are recorded in conjunction with other electrical activity, deriving from the brain and the muscles of the face, scalp and neck. Our measurements consider the steady-state responses as 'signal' and the other activity as 'noise'. An assessment of the signal-to-noise ratio and an estimate of whether the signal is significantly different from noise are essential for any meaningful interpretation of the response. Different tests for determining whether a recorded response is significantly different from noise are illustrated in Figure 3.

One approach considers the variance of repeated measurements of the response. For example, one can record 16 separate responses, each over a period of 1.024 s. A second approach considers the variance of the different amplitudes in a single spectrum. For example, one can join together the 16 responses into one 16.384-s sweep and perform an FFT to give a spectrum of amplitudes.

*Repeated measurements*
Since the response is two-dimensional, the variance is likewise two-dimensional. Hotelling's $T^2$ test, a multivariate analog of the univariate $t$-test (Hotelling, 1931; Anderson, 1984), can show the confidence limits of a mean two-dimensional response

using an ellipse. If the ellipse does not include the origin, the probability that the response equals zero is less than the probability of the confidence limits (Picton, 1987; Picton et al, 1987a). Given a set of $N$ paired measurements $x_j$ and $y_j$ with means and and covariance matrix $S$, one can calculate a statistic $T^2$ to determine whether the measured means are significantly different from zero:
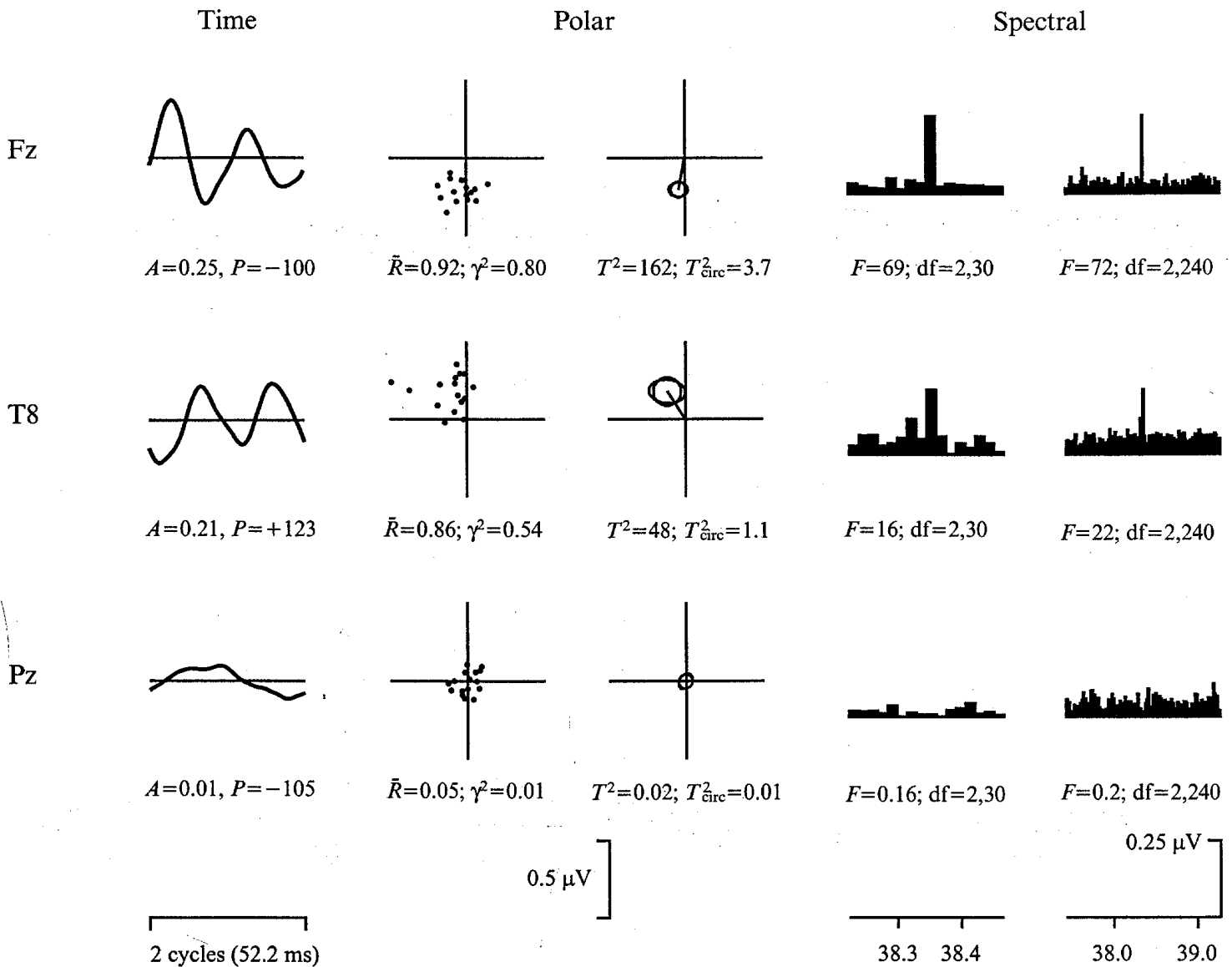
$$T^2 = N[\bar{x},\bar{y}]' S^{-1}[\bar{x},\bar{y}]$$

Multiplying $T^2$ by $(N-2)/(2N-2)$ gives a value that is distributed as $F$ with 2 and $N$–2 degrees of freedom. Graphically, one can plot the confidence limits of the mean by creating an ellipse using the value of $F$ required for the degrees of freedom and significance level ($\alpha$). If the 0,0 point is outside the ellipse, then the probability is less than $\alpha$ that the measured mean is zero.

Victor and Mast (1991) assumed equivalent variance in the real and imaginary dimensions, and proposed a 'circular $T^2$ test' as a simpler and more powerful estimate of the confidence limits:

$$T_{circ}^2 = (N-1)(\bar{x}^2 + \bar{y}^2) / \sum_{j=1}^{N}\left[(x_j - \bar{x})^2 + (y_j - \bar{y})^2\right]$$

Multiplying $T_{circ}^2$ by $N$ gives a value that is distributed as $F$ with 2 and $2N$–2 degrees of freedom. The confidence limits for the mean are then plotted as a circle instead of an ellipse.

|  | Time | Polar | | Spectral | |

**Fz**

$A=0.25$, $P=-100$    $\bar{R}=0.92$; $\gamma^2=0.80$    $T^2=162$; $T^2_{circ}=3.7$    $F=69$; df=2,30    $F=72$; df=2,240

**T8**

$A=0.21$, $P=+123$    $\bar{R}=0.86$; $\gamma^2=0.54$    $T^2=48$; $T^2_{circ}=1.1$    $F=16$; df=2,30    $F=22$; df=2,240

**Pz**

$A=0.01$, $P=-105$    $\bar{R}=0.05$; $\gamma^2=0.01$    $T^2=0.02$; $T^2_{circ}=0.01$    $F=0.16$; df=2,30    $F=0.2$; df=2,240

0.5 μV

0.25 μV

2 cycles (52.2 ms)    38.3   38.4    38.0    39.0

**Figure 3.** Signals in noise. The data for this figure were obtained in the same way as those used in Figure 1, but from a different subject. Responses are shown from three different scalp locations: midfrontal (Fz), right temporal (T8), and midparietal (Pz) (see Figure 14 for locations). On the left, the responses are plotted as an average time waveform over two cycles of stimulation. Below the waveforms are shown the amplitude in μV and the cosine phase at the onset of the waveform in degrees. The next column shows 16 replications of the response (from each 4 s) plotted in polar form. Below these plots are shown the phase coherence ($\bar{R}$) and the magnitude squared coherence ($\gamma^2$). The next column shows the $p<0.05$ confidence limits of the mean for the data plotted in the second column, assuming independent (ellipses) or equal (circles) real and imaginary variance. Below these plots are given the $T^2$ and $T^2_{circ}$ measurements. These can be converted to $F$-values by using the equations in the text. The right two columns show spectral plots of the response, using a range of 16 or 120 frequency bins. Below each are given the $F$-values for comparing the amplitude at the response frequency to the amplitude in the adjacent frequencies. The T8 recording shows a significant response that has a different phase from that at Fz. The response at Pz is not significantly different from zero.

A steady-state response might hypothetically vary separately in amplitude and phase (Picton et al, 1987b). For example, phase might be totally determined by the intensity of the stimulus, while amplitude might vary with both the intensity of the stimulus and the state of arousal. In such cases, the original $T^2$ statistic is a more appropriate measurement. Furthermore, the power advantage of the circular $T^2$ over the original $T^2$ statistic decreases as the number of responses increases (Valdes et al, 1997).

A measurement of the 'magnitude squared coherence' (Dobie and Wilson, 1989, 1995a) calculates $\gamma^2$ as the power of the average response divided by the average power of the individual responses:

$$\gamma^2 = N\left(\bar{x}^2 + \bar{y}^2\right) / \sum_{j=1}^{N}\left(x_j^2 + y_j^2\right)$$

The result is between 0 and 1, with higher values meaning a higher signal-to-noise ratio and a lower probability that the response equals zero. The preceding two formulae are very similar, and it can be shown that magnitude squared coherence and the circular $T^2$ test are mathematically equivalent (Dobie and Wilson, 1993).

The variance of the repeated measurements can also be considered simply with respect to phase. Phase coherence (Jerger et al, 1986; Stapells et al, 1987) is a value that derives from earlier measurements of phase variance (Mardia, 1972; Sayers et al, 1979; Fridman et al, 1984; Aoyagi et al, 1993a). A value ($\bar{R}$) is calculated as the root mean square of the sums of the cosines and sines of the individual phase angles:

$$\bar{R} = \left[\left(\sum_{j=1}^{N}\cos\theta_j\right)^2 + \left(\sum_{j=1}^{N}\sin\theta_j\right)^2\right]^{0.5} / N$$

The Rayleigh test (Rayleigh, 1880) assesses the probability that this measurement could result from a random sampling of all possible phases. The significance of the result can be estimated by using simple formulae (Fisher, 1993). Essentially, this measurement is the same as magnitude squared coherence with the amplitude removed.

A statistic combining both amplitude and phase should be more powerful than one using only phase, and is demonstrably so for simulated data (Dobie and Wilson, 1993). However, several comparisons using actual recordings of steady-state responses have shown that phase coherence is not significantly different from magnitude squared coherence or Hotelling's $T^2$ when trying to detect the presence or absence of a response (Picton et al, 1987a; Champlin, 1992; Dobie and Wilson, 1994a; Stürzebecher and Cebulla, 1997; Valdes et al, 1997). Recent studies have shown a small but clear advantage for measurements combining amplitude and phase (Picton et al, 2001; Cebulla et al, 2001).

## Measurements of the spectrum

The level of the background noise in a recording can be estimated by measuring the activity at frequencies in the spectrum other than that of the stimulus and response. Comparing the power of the signal to the powers at other frequencies is the basis of the F-test for hidden periodicity (Schuster, 1898; Fisher, 1929; Zurek, 1992; Dobie and Wilson, 1996; Lins et al, 1996). The procedure calculates an $F$ ratio of the power in the signal frequency bin ($s$) to the mean power in $N$ adjacent bins:

$$F = N\left(x_s^2 + y_s^2\right) / \sum_{\substack{j=s-N/2 \\ j\neq s}}^{s+N/2}\left(x_j^2 + y_j^2\right)$$

This is distributed as $F$ with degrees of freedom 2 and $2N$. The F-test will clearly vary with the number of adjacent frequencies used in the comparison. A higher number will increase the power of the test by increasing the denominator degrees of freedom. However, extending the range of the noise samples too

far is counteracted by the fact that the background EEG noise is not white, but shows decreasing power with increasing frequency. We have found that using 16 noise samples produces similar results to using 120 samples (cf Dobie and Wilson, 1991, 1996). The F-test is essentially the same as the magnitude squared coherence when the number of individual measurements for calculating the coherence equals one less than the number of adjacent points used in the F-test, i.e. the degrees of freedom are the same (Dobie and Wilson, 1996).

The F-test has several advantages over tests based on repeated measurements of the response. First, the number of adjacent frequency bins to which the signal response is compared can be increased beyond any easily obtained number of separate measurements of the signal response. Second, the technique can easily be adapted to omit certain frequency bins from the calculation. In this way, a noise estimate can be obtained that is uncontaminated by line noise or by responses at other frequencies (if one is recording responses to multiple simultaneous stimuli).

## Averaging

Both averaging and filtering can increase the signal-to-noise ratio of a recording. If a signal is time-locked to the recording sweeps (and the noise is not), averaging multiple sweeps together will maintain the signal but decrease the amplitude of the noise. Filtering is the process by which a particular frequency (or set of frequencies) is maintained while other frequencies are attenuated. Choosing to consider just one frequency in the spectrum provided by the FFT is effectively using a bandpass filter with a width equal to the resolution of the FFT (the reciprocal of the sweep duration). This filter can be made more exact by increasing the duration of the recording sweep that is submitted to the FFT. Given a long sweep of recorded data, one can average the data over fractions of the sweep to reduce the noise. For example, 16 s of data can be divided into 16 separate 1-s sweeps, and these can be averaged together and submitted to an FFT with a resolution of 1 Hz. On the other hand, one can also submit the whole 16-s sweep to an FFT with a resolution of 0.0625 Hz. Both procedures work equally well to reduce the noise (Regan, 1989; John et al, 1998; Pethe et al, 2000). When measuring the auditory steady-state responses, one customarily uses both averaging and the FFT 'filter' to reduce the noise. The length of the sweep submitted to the FFT can be determined by the speed and memory of the computer, and by the resolution required to discriminate the response. For example, if one is recording two simultaneous responses at 80 and 81 Hz, one should use a sweep of longer than 1 s in order to separate the two responses in the spectrum. If there are an integer number of cycles in the sweep the steady-state response will fall discretely into a single FFT bin, and the FFT resolution can be made extremely precise if the sweep is very long (Pethe et al, 2000).

Brief periods with increased noise levels often occur during a recording, due to events such as movement or swallowing. These periods may be removed from the averaging process through a process of artefact rejection, based on the maximum amplitude of the input (or some other amplitude measurement). Two considerations are important when setting up artefact rejection for steady-state responses. First, the data might be filtered to pass only frequencies near those of the steady-state response, before

determining whether to reject. For example, when measuring the 80-Hz steady-state responses, one might wish to reject trials on the basis of noise at frequencies near 80 Hz and not reject trials because the subject's alpha rhythm at 10 Hz exceeds the rejection criteria. Second, when using a long sweep, one should not reject the whole sweep on the basis of a brief period of high noise. If possible, the long sweep should be broken up into smaller epochs, and only the noisy epoch rejected. Subsequently recorded data can be used to continue the sweep in place of the rejected epoch (John and Picton, 2000b). This technique requires each epoch to contain an integer number of stimulus cycles and to start at the same stimulus phase.

Another approach to the problem of intermittent noise is to use weighted averaging. Each recorded epoch is divided by its variance before being summed with the other epochs. The final sum is then divided by the sum of the reciprocals of these variances. High-noise epochs thus have little effect on the final result. This procedure works well for transient evoked potentials (Hoke et al, 1984; Lütkenhöner et al, 1985; Elberling and Wahlgreen, 1985) and for simulated steady-state responses (Dobie and Wilson, 1993). We have found it very effective in enhancing the detectability of near-threshold steady-state responses (John et al, 2001a). The variance estimates should be based on a signal that is bandpass-filtered in the region of the response, and the variance weighting of the recorded signal should be applied using brief recording epochs (of the order of 1 s). Figure 4 shows the use of normal and weighted averaging in recording multiple auditory steady-state responses.

*Multichannel recordings*
Most recordings of the auditory steady-state responses have been made from a single 'channel', usually between the vertex or midfrontal electrode and a reference on the mastoid or neck. Recording from multiple channels might be helpful in two ways. First, these recordings can be analyzed to provide some estimate of the intracranial sources of the responses. The response is generated in one or several regions of the brain, each of which generates electrical fields that can be recorded at the scalp. Each generator has a specific transfer function for each region of the scalp, so that the distribution of the evoked potential over the scalp (its 'scalp topography') is quite specific to the generator. Source analysis procedures can estimate the sources from the scalp-recorded potentials, provided that the transfer functions between source and scalp are reasonably well known (see 'Source analysis' below).

Second, the detection of the responses might be facilitated by examining the recordings from multiple channels rather than from just one. Multiple channels of data recorded from the scalp can be analyzed together, provided that attention is paid to the covariance between these channels. The F-test, described in 'Measurements of the spectrum' above, can be expanded to include data from multiple channels. Inasmuch as the noise is uncorrelated between channels and the signal present in each channel, a multichannel assessment of whether a signal is present may be more powerful than any single-channel assessment. However, the fact that the signal may be very small or the noise very large in particular channels may decrease the power of a multichannel analysis.

Pedro Valdes-Sosa and the first author extended the single-channel F-test to the multichannel condition in the following way.

A complex vector $z$ is constructed from each of the FFTs performed separately on each of $K$ channels, of length equal to one-half the size of the data vector. From these FFT results, $J$ data points are taken to represent the noise. Only those frequencies near the signal frequency are used. A matrix $Z$ of complex values ($K$ channels in rows; $J$ FFT data in columns) is then constructed from the $z$ vectors for each of the channels. A covariance matrix $S$ is calculated for the noise data across the channels:

$$S = ZZ'/J$$

The signal vector $a$ is set up as a complex vector with a length equal to the number of channels. The probability of the signal being larger than the noise is then estimated from the result of:

$$2aS^{-1}a'$$

This value is distributed as Hotelling's $T^2$, which is related to $F$ according to the equation:

$$F_{2K,2J-2K+2} = T^2(2J-2K+2) / [(2J)(2K)]$$

Our formulations were derived from the work on complex normal distributions by Brillinger (1975), particularly exercise 4.8.28, p. 113. The formulations are similar to those for non-complex data (e.g. Mardia et al, 1979, p. 74), except for the setup of the covariance matrix with the conjugation of the imaginary components (accounting for the independence of the real and imaginary components), and the doubling of many of the parameters.

The usual clinical approach is to record from the one channel where the signal-to-noise ratio is presumed to be close to optimal. We compared this approach with the multichannel analysis for auditory steady-state responses to amplitude modulation with modulation frequencies near 80 Hz recorded from 47 scalp channels. The multichannel F-test was not more powerful than the single-channel F-test on the channel (vertex to neck) with the highest signal-to-noise ratio. If this channel is known beforehand, the multichannel F-test does not significantly increase our ability to detect the response. The 80-Hz responses have a simple scalp topography, since they are predominantly generated by a single brainstem source. The multichannel $F$ may be more effective if the source generators and scalp topography are more complicated.

If one knows beforehand the intracranial source of the response, one can attenuate the noise recorded from multiple scalp channels by projecting both the signal and the noise to the generator site. The basic idea of this 'source projection' technique is that the noise recorded in the scalp channels is generated at other locations in the brain as well as at the site of response generation. Figure 5 shows the spectrum of the responses at the source location compared to the spectrum of the responses in the vertex–neck recording. Our initial work with this approach has shown that source projection can increase the signal-to-noise ratio a little over the optimal scalp recording, but that this increase may not be sufficient to warrant the extra time required to position the multiple electrodes and perform the extra analyses. Again, if the optimal channel is known beforehand, this single channel will be more efficient than a multichannel recording.
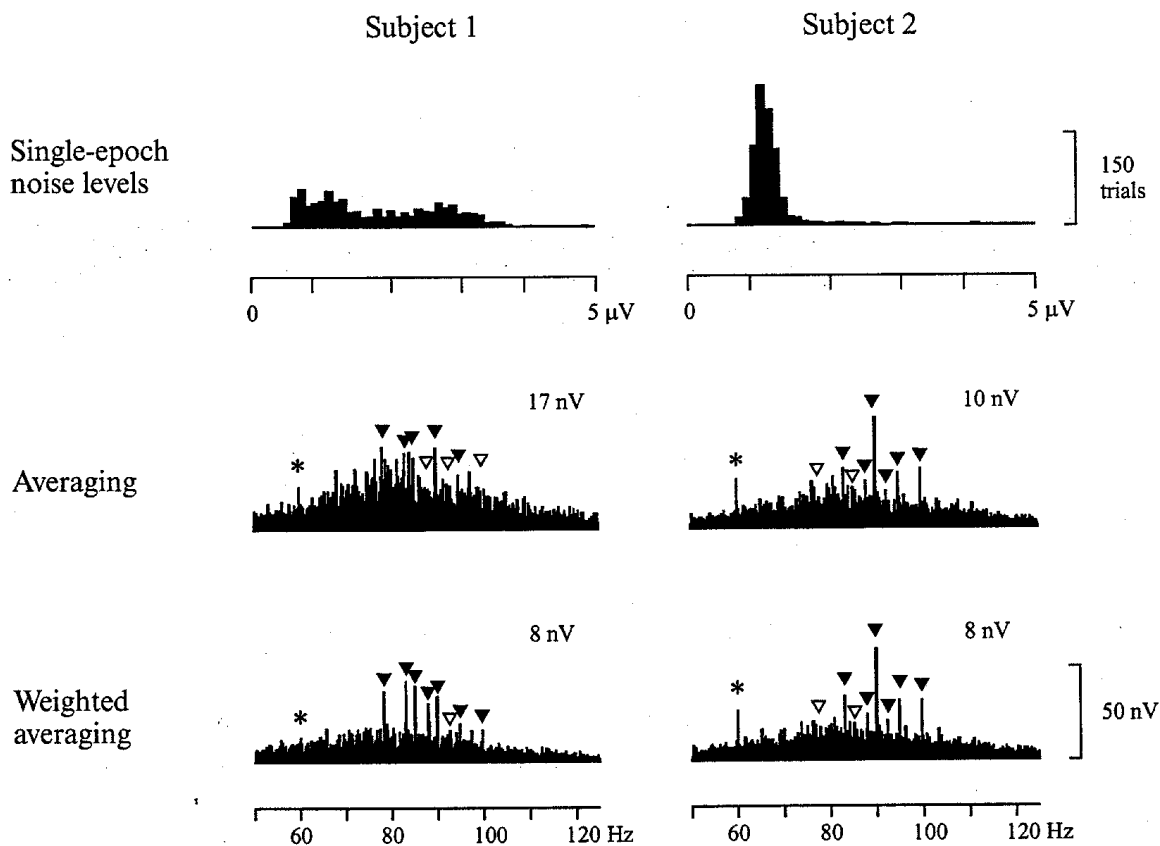
**Figure 4.** Weighted averaging. This figure shows the auditory steady-state responses recorded using eight simultaneous tones modulated at frequencies of 78.1, 83.0, 85.0, 87.9, 89.8, 92.7, 94.7 and 99.6 Hz. The recording was bandpass-filtered between 65 and 110 Hz. The final analysis was based on a 16.384-s sweep made up of 16 epochs of 1.024 s. Averaging was carried out over 48 sweeps prior to the analysis. For weighted averaging, each epoch was weighted by dividing it by its variance. For the first subject, the variance varies significantly from epoch to epoch, as shown by the histogram of the standard deviations of the amplitudes within an epoch, plotted at the top of the figure (based on 768 epochs). In this subject, weighted averaging significantly increases the signal-to-noise ratio, the final noise level decreasing from 17 to 8 nV (numbers to the upper right of each response). The responses are indicated by the triangles, with the filled triangles indicating responses that reached a significance level of $p < 0.01$, and open triangles indicating responses that were not significantly different from noise. Weighted averaging makes it possible to recognize seven rather than five responses. The second subject does not show much variability in the noise levels within single epochs, and weighted averaging causes little change in the noise levels and no change in the number of responses recognized (six). The asterisks indicate line noise at 60 Hz.

## Biasing detection

When looking for a signal in a haystack of noise, it is helpful to know what the signal looks like. Dobie and Wilson (1994b) proposed using phase weighting to bias the detection of a response. Responses that are closer in phase to an expected phase are considered to be more significant. The expected phase can be derived from normative data or from previous responses recorded in the same individual. For example, one can use the phase delay of the response at high intensity to predict the phase delay of the response at lower intensity (which should be a little longer). Lins et al (1996) attempted to bias the results by how closely both the amplitude and phase of a response fit with the mean response from a group of normal subjects. However, this was not as successful as simply using phase, perhaps because there are larger intersubject differences in amplitude than in phase. The difficulty with these approaches is that significance levels for testing can only be estimated after modeling the effects of

biasing. One approach, suggested to the first author by Robert Dobie, would be to project the two-dimensional data onto a one-dimensional axis oriented at the expected phase. The confidence limits for these one-dimensional data can then be assessed using a simple $t$-test, as illustrated with modeled data in Figure 6. We have adapted this idea to the F-test, and shown that phase weighting the evaluation can lead to better recognition of responses, provided that there is some reliable *a priori* indicator of the expected phase (Picton et al, 2001).

## Combining information

The steady-state response often contains significant energy at more than one harmonic of the stimulus frequency. How does one combine the information at these different frequencies to detect the response? A similar situation occurs when multiple different responses are recorded simultaneously, e.g. responses to one carrier frequency modulated by several different
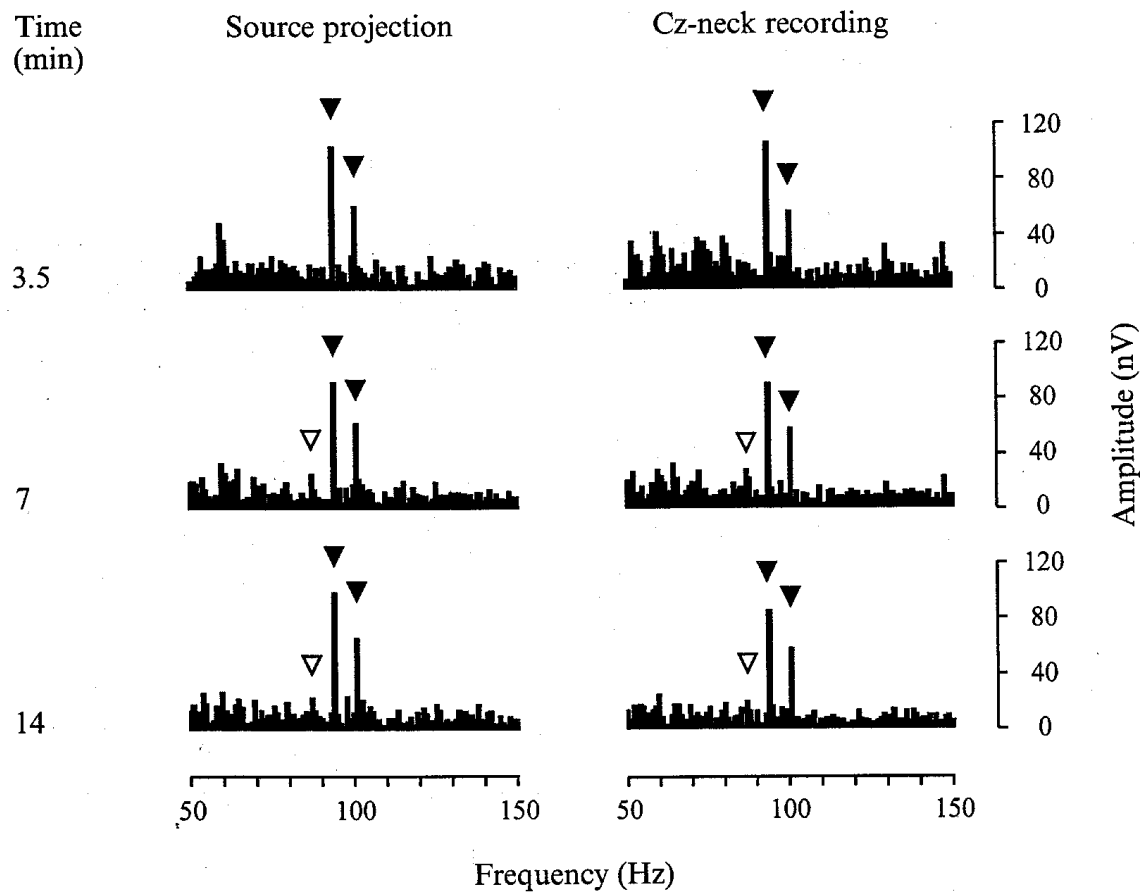
**Figure 5.** Multichannel evaluation of signal in noise. The data for this figure were collected from 47 scalp electrodes. Three simultaneous AM stimuli were presented binaurally at 40 dB SPL. The carrier frequencies were 500, 1000 and 2000 Hz, and the modulation frequencies 86.9, 93.8 and 100.6 Hz, respectively. Sweeps of 1.024 s were averaged 200, 400 and 800 times. The spectra in the right column show the responses recorded in the Cz-neck channel. The responses for the 1000-Hz and 2000-Hz carriers were significant at $p<0.01$ (filled triangles) after the first 200 sweeps. The response to the 500-Hz carrier became significant at $p<0.05$ (open triangles) after 400 sweeps. The spectra in the left column are those obtained after projecting the scalp recordings onto a dipole generator located in the center of the head and oriented upward and 15° anteriorly. This response shows similar signal-to-noise levels to the Cz-neck recording. The amplitude of the source response is arbitrary (depending on the estimated transfer characteristics of the head), and has been adjusted to be similar to that of the scalp recording.

modulation frequencies. If the main question concerns the presence or absence of a response at the carrier frequency, how does one combine the information across the different responses? Stürzebecher et al (1999) have provided some modifications of phase coherence and magnitude squared coherence for $q$ spectral samples. Dimitrijevic et al (2001) applied the Stouffer test, normally used in meta-analysis, to combine the probabilities of two separate responses involving the same carrier frequency. Two independent F-tests provided probability levels for each response, which were then combined to yield the probability of a response being evoked at that particular carrier frequency.
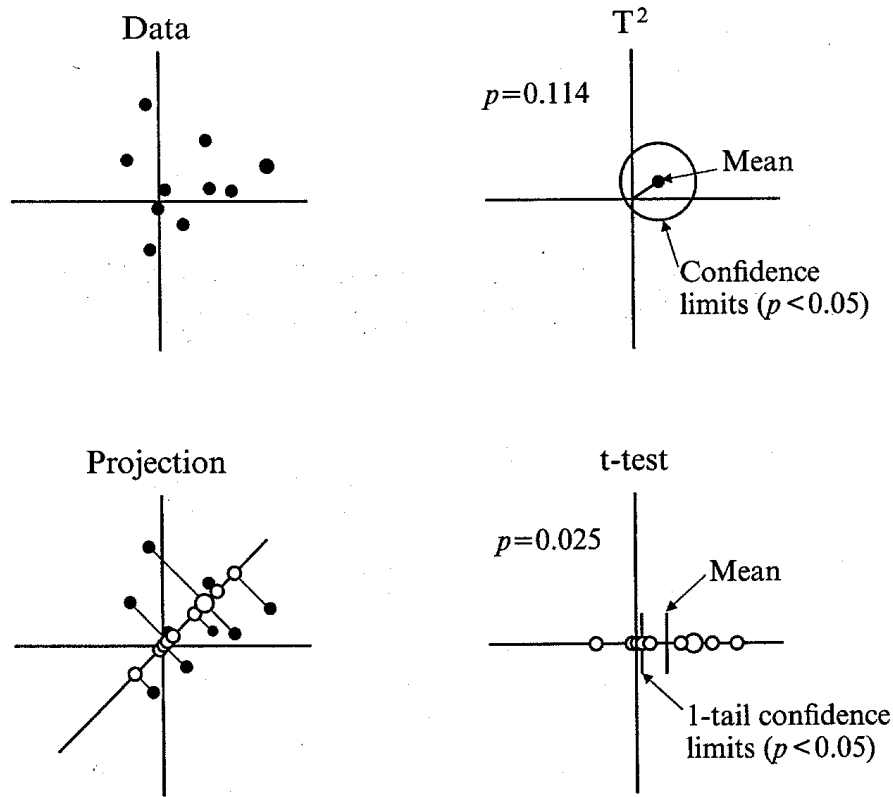
*Characteristics of the EEG noise*

In our present approaches to detecting the steady-state response, very little attention is paid to the characteristics of background noise. The electrical activity of the brain is called the electroencephalogram or EEG. Averaging assumes that the EEG is random in relation to the signal, and the F-test assumes that it is equal across the frequencies adjacent to the signal frequency. Neither assumption is true. The EEG is often organized into
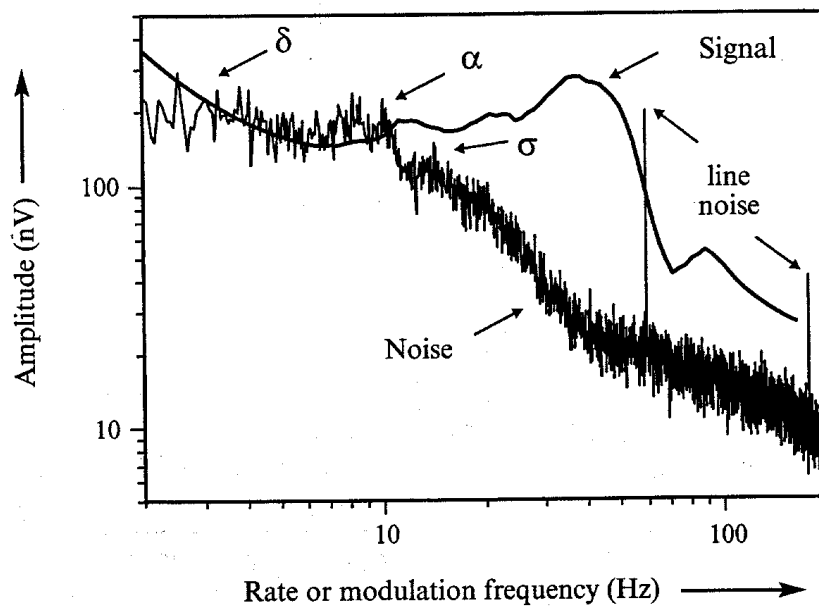
rhythms that do not behave randomly, and the amount of energy in the EEG spectrum decreases with increasing frequency. When recorded from the scalp, the EEG is itself intermixed with other electrical activity coming from the scalp muscles, the eyes, the skin, and the tongue.

The FFT is very effective in recognizing responses at a particular frequency. However, when the response frequency is not exactly the same as the frequency of one of the bins in the FFT (that is, the frequency is not an integer multiple of the resolution), there will be some 'leakage' of its energy into other bins. We can take care of this from the point of view of the response by presenting an integer number of modulation cycles within the sweep. However, this does not prevent noise energy from leaking into the frequency bin of the stimulus. This leakage can be attenuated by filtering the signal prior to the FFT. Tang and Norcia (1993) have shown that this improves the recognition of the visual steady-state response, but this has not yet been done for the auditory responses.

Figure 7 shows the spectrum of the scalp-recorded activity averaged across a group of drowsy or sleeping subjects during experiments evaluating the steady-state responses (Picton et al,

Data

T$^2$

$p=0.114$

Mean

Confidence
limits $(p<0.05)$

Projection

t-test

$p=0.025$

Mean

1-tail confidence
limits $(p<0.05)$

**Figure 6.** Phase-weighting. This figure illustrates the idea of phase-weighting using the modeled data shown in the upper left. The $T^2_{circ}$ measurement shown on the upper right indicates that the zero point is included within the two-dimensional confidence limits of the mean of the data. The response is therefore not significantly different from zero. At the lower left of the figure, the data are projected onto the line of the expected phase of the response. These projected data can then be evaluated to see whether the confidence limits of the mean (obtained using a one-tailed $t$-test) include the zero point. In this example, the additional information available from the expected phase allows us to recognize a significant response that is not significant if we do not consider the phase.

$\delta$

$\alpha$

Signal

$\sigma$

line
noise

100

Amplitude (nV)

Noise

10

10

100

Rate or modulation frequency (Hz)

**Figure 7.** Measurements of signal and noise at different frequencies. The noise data were obtained from the spectra recorded when no stimuli were presented (Picton et al, 2001). For each of 20 subjects, 24 sweeps of 16.4 s were combined using weighted averaging. The plotted data represent the average (phase disregarded) of the amplitude spectra across the 20 subjects. The two vertical lines represent the activities at the frequency of line noise and its third harmonic (60 and 180 Hz). The subjects were drowsy or asleep, and the background EEG noise shows slow delta activity ($\delta$) with small peaks at the frequency of the alpha ($\alpha$) rhythm (10 Hz) and spindle ($\sigma$) activity (14–16 Hz). The smooth line superimposed on the noise data plots the general amplitudes of the steady-state responses from Figure 10 using log–log coordinates rather than the semilog coordinates used in that figure.

2001). The EEG was recorded under control conditions with no stimuli. The amount of activity decreases with increasing frequency, with the activity being most prominent at frequencies below 25 Hz. At higher frequencies, the noise decreases exponentially with increasing frequency (Cohen et al, 1991). Because of this, steady-state responses at high frequencies may be easily recognizable despite being smaller than responses at low frequencies. The EEG spectrum can be modeled with a basic pink noise having a power that decreases exponentially with frequency, and several superimposed peaks where the alpha and beta rhythms occur (Isaksson and Wennberg, 1975; Dumermuth and Molinari, 1987). The amount of slow activity increases with slow-wave sleep, anesthesia, and brain damage. Increased tension in the muscles underlying the scalp can increase the activity in a broad range of frequencies. This activity has not been extensively studied, but appears to have its maximal energy between 10 and 100 Hz.

The EEG noise may show significant relationships between frequencies and across different recording locations. The noise may be much more organized than we have thought, and we may be able to predict the noise from one moment to the next. If so, we might be able to model the noise as distinct from the signal, and reject it from our recordings far more efficiently than by averaging. Tang and Norcia (1995) have shown that this type of adaptive filtering can significantly improve the detection of the visual steady-state response over both the simple FFT and a combination of FFT and prefiltering.

## Stimulus relations

### Types of stimuli

Figure 8 shows examples of the various kinds of stimuli that have been used to evoke auditory steady-state responses. The first auditory steady-state responses were evoked by clicks and tone-bursts. These stimuli have energy at multiple frequencies in the acoustic spectrum. Campbell et al (1977) recorded the first auditory steady-state responses to sinusoidally amplitude-modulated (AM) tones. These stimuli have energy at the frequency of the carrier and at two sidebands separated from the carrier by the modulation frequency. Another kind of stimulus is the 'beating' that occurs when two tones are presented together. The beating occurs at the difference between the frequencies of the two component tones. These two-tone stimuli evoke similar responses to those evoked by sinusoidally AM tones in gerbils (Dolphin et al, 1994), whales (Dolphin et al, 1995), dolphins (Popov et al, 1996), and chinchillas (Arnold and Burkard, 1997).

AM noise can also be used to elicit steady-state responses. If the noise spans a range of frequencies, it will activate auditory fibers over a broad region of the basilar membrane. The compound response of all these fibers is larger than the response to a pure tone (John et al, 1998). However, the effect of an increased number of responding fibers is partially counteracted by the phase dispersion caused by the differences in timing from one region of the basilar membrane to another (as determined by the traveling wave and the filter characteristics of the hair cells). The advantage of using noise is the greater amplitude of the responses, which decreases the time needed to demonstrate a significant response. AM white noise might thus be helpful as a rapid screening test for hearing in much the same way as the click (which also has a broad frequency spectrum) used to evoke

auditory brainstem responses. The disadvantage of the noise is the lack of frequency specificity. However, it is possible that decreased hearing at particular frequencies might significantly change the amplitude or the phase of the response. Figure 9 shows steady-state responses to AM white noise.

### Stimulus rate

One of the most defining characteristics of the auditory steady-state responses is their relationship with the rate at which stimuli are presented. Some of the responses, such as the 40-Hz evoked potential, derive their names from this stimulus rate. The effect of rate was examined in most of the initial studies of the steady-state responses. Geisler (1960) found a general decrease in the peak-to-peak amplitude of the response as click rate was increased from 1/s to 120/s. Campbell et al (1977) recorded the responses to clicks and AM tones at rates between 6 and 32 Hz, and found larger responses at the middle rates, although for the 500-Hz AM tone the response was larger at 32 Hz. Galambos et al (1981) showed significant enhancement of the response at rates near 40 Hz, and these results were confirmed by Stapells et al (1984). Rickards and Clark (1984) recorded steady-state responses to AM tones with modulation frequencies between 4 and 450 Hz. The amplitude of the responses decreased with increasing modulation frequency, with some enhancement of the amplitude at frequencies near 40 Hz. Rees et al (1986) recorded steady-state responses to AM tones using modulation frequencies between 20 and 400 Hz. The modulation transfer function for their steady-state responses showed the characteristics of a low-pass filter with a cut-off of 40–50 Hz and some additional enhancement near 40 Hz. Herdman et al (2002a) found that the response to a 70-dB SPL 1000-Hz tone was five times larger when the stimulus was modulated at 39 Hz than when it was modulated at 88 Hz, when recorded in waking subjects.

Figure 10 combines data from these and other studies. Two basic principles govern the effects of rate. The first is that the responses generally decrease with increasing stimulus rate. The second is that the response is enhanced above this general decline in certain regions, most clearly near 40 Hz, but also near 90 Hz. As discussed in 'Characteristics of the EEG noise' above, the noise in which the responses are recorded also decreases with increasing frequency. Therefore, although the response amplitude decreases with increasing modulation frequency, the signal-to-noise ratio may actually increase.

Because the background EEG noise decreases with increasing frequency, some measure of the detectability of the response may be used in addition to amplitude to assess the modulation transfer function. Picton et al (1987c) found that the noise levels at the low frequencies (<20 Hz) made responses (to either AM or FM) very difficult to recognize using the $T^2$ test. Levi et al (1993) measured the magnitude squared coherence of the response between 10 and 80 Hz, and found that, although the 40-Hz response had higher values during wakefulness, the 80-Hz response showed higher values during sleep. Dobie and Wilson (1998) showed two regions of enhanced detectability—at 40–50 Hz and at 80–100 Hz. They replicated the results of Levi et al (1993) at high intensity levels, but found that, at low intensity levels, the 40-Hz response remained more detectable in both waking and sleeping subjects.

Linden et al (1985) reported that the auditory steady-state response to tone-bursts presented at rates between 10 and 60 Hz
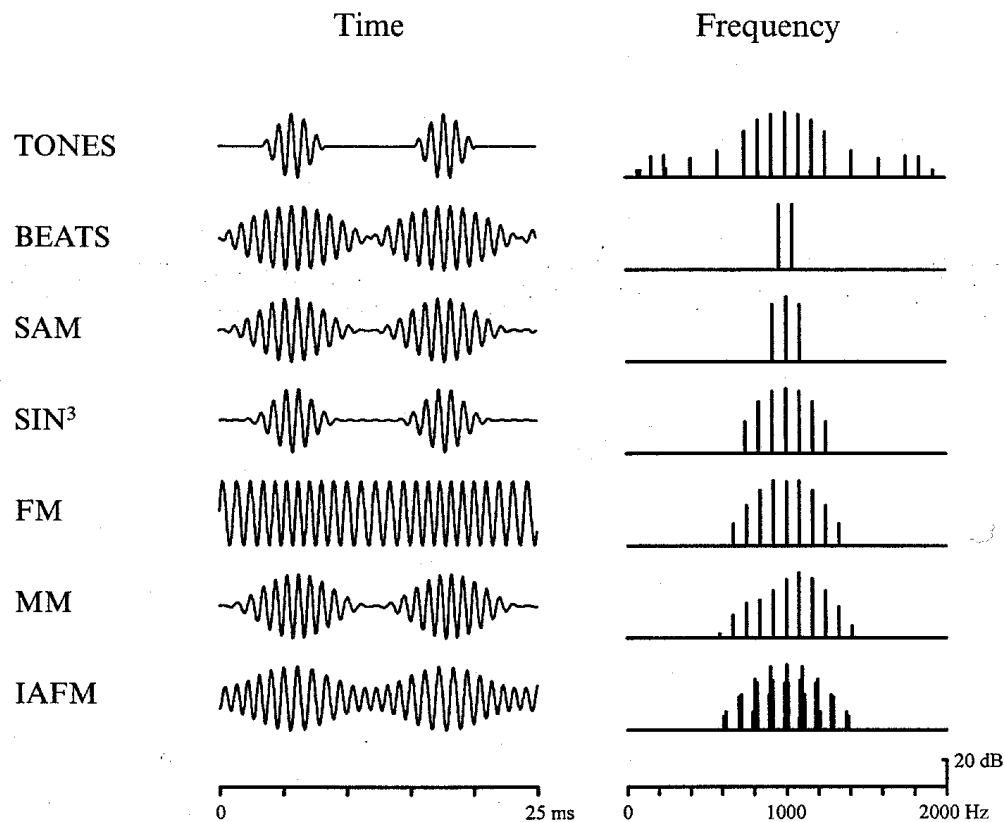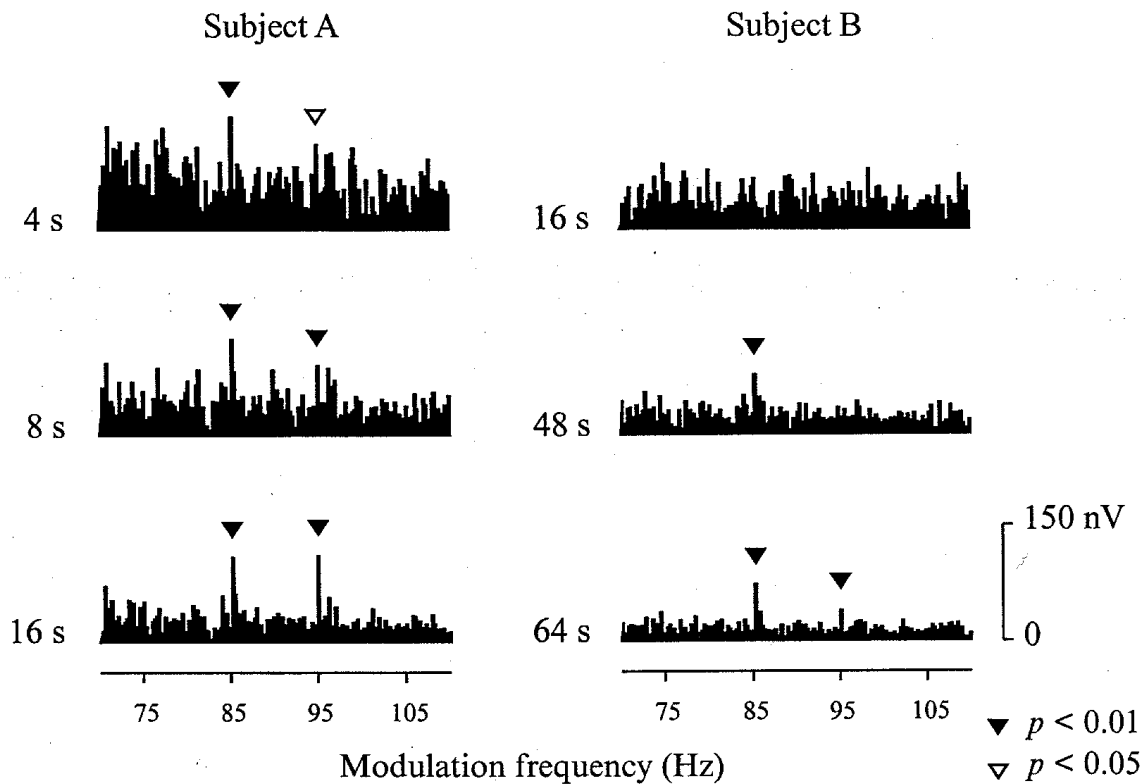
**Figure 8.** Stimuli used to evoke auditory steady-state responses. This figure shows various types of stimuli that have been used to evoke the auditory steady-state responses. The stimulus waveforms are plotted in the time domain on the left, and the spectra of the stimuli (based on a much longer time sample) are shown on the right. These data were obtained by calculation; electric and acoustic waveforms would be basically similar, with some filtering effects during passage through the signal generators and transducers. The TONES represent brief tone-bursts of 1000 Hz with the commonly used 2–1–2 envelope, with rise and fall times of two cycles (2 ms) and a plateau of 1 cycle (1 ms). The spectrum shows a broad splatter of energy into frequencies far removed from the nominal frequency of the tone. The BEATS were obtained by adding together continuous tones of 958 and 1042 Hz. The sinusoidally amplitude-modulated (SAM) tone used a carrier of 1000 Hz and a modulation frequency of 84 Hz. There are spectral peaks at the carrier frequency and at one sideband above and below the carrier, separated from it by the modulation frequency. The $SIN^3$ tone used a modulation envelope based on the third power of the usual sinusoidal envelope (John et al, 2002a). The spectrum contains three sidebands on either side of the carrier. The FM tone is sinusoidally modulated with a depth of modulation of 25%. The MM tone has mixed modulation (100% AM and 25% FM), with both modulations occurring at 84 Hz. The bottom set of data represent independent amplitude and frequency modulation or IAFM, with 50% amplitude modulation at 84 Hz and 25% frequency modulation at 98 Hz. The spectrum shows a complex set of peaks, but all the energy remains concentrated around 1000 Hz. All the time waveforms are plotted so that they have the same peak amplitude. The spectra are plotted logarithmically over a range of 50 dB relative to the maximum peak in the spectrum.

showed a maximum amplitude between 30 and 50 Hz. This maximum occurred at lower rates during sleep, but the differences were not significant. Sapsford et al (1996) (see also Andrade et al (1996) and Munglani et al (1993)) have measured the auditory steady-state responses to clicks presented at rates between 6.5 and 50.5 Hz, and calculate a coherence index as the sum of the power at the fundamental frequency of stimulation and its harmonics (up to the fourth). A third-order polynomial trend is fitted to the data, and the rate at which this trend line reaches its maximum value is taken as a measure of cognitive function. In wakefulness, the maximum coherence occurs within 37 and 40 Hz (95% confidence limits). This value decreased significantly when low doses of propofol or isoflurane were administered, and this decrease varied with the decrease in performance on a simple memory test.

A general idea is that the responses to the higher rates are mediated by lower stations of the auditory system. This is probably an oversimplification. Sequential low-pass filtering occurs, with the lower stations responding to a wide range of frequencies, and the higher stations responding to a decreasing range of frequencies. Thus, responses at modulation frequencies near 80 Hz are probably generated mainly in the brainstem, whereas the responses at modulation frequencies near 40 Hz are generated in both the brainstem and cortex.

*Amplitude modulation*
The most commonly utilized steady-state stimuli are sinusoidally AM tones. These stimuli contain spectral energy at the carrier frequency and at two sidebands on either side of the carrier, at a frequency separation equal to the modulation frequency. When

Subject A  Subject B

4 s

16 s

8 s

48 s

150 nV

16 s

64 s

0

75    85    95    105          75    85    95    105       ▼ $p < 0.01$

Modulation frequency (Hz)                                  ▽ $p < 0.05$

**Figure 9.** Responses evoked by AM noise. This figure shows the steady-state responses to AM noise recorded from two subjects. The stimulus was a continuous white noise modulated using an exponential envelope (see 'Amplitude modulation') with a power of 2. The intensity of the stimulus was 50 dB rms SPL, which was approximately 30 dB above the average behavioral threshold for this stimulus. Two stimuli were presented simultaneously, one to each ear, with a modulation frequency of 85 Hz for the left ear and 95 Hz for the right ear. Responses were checked at integer multiples of the sweep, which, for this particular set of recordings, was set at 4.096 s. The responses were analyzed using an F-test, and responses considered significantly different from noise are indicated with arrowheads (open at $p<0.05$ and closed at $p<0.01$). The response amplitudes are larger than those seen with modulated tones (Figure 4). For subject A, both responses were recognized on the first sweep (4 s). For subject B, the responses were smaller and the background EEG a little higher (compare the responses after 16 s of averaging). The responses for this subject were not recognized until after 48 s of averaging for the left ear and 64 s for the right ear. In our first group of 10 adult subjects, 75% of responses were recognized within 1 min, and all responses were recognized within 3 min.

the modulation is 100%, the amplitude at the sidebands is one-half that at the carrier frequency. The formula for a digitized AM tone is:

$$s(i) = a[1 + m_a \sin(2\pi f_{am} t i)] \sin(2\pi f_c t i)$$

where $i$ is the address in the output buffer, $a$ is the amplitude of the stimulus, $t$ is the time per address at which digital-to-analog (DA) conversion occurs, $f_{am}$ is the modulation frequency for amplitude modulation, and $f_c$ is the carrier frequency. The depth of amplitude modulation $(m_a)$ is defined as the ratio of the difference between the maximum and minimum amplitudes of the signal to the sum of the maximum and minimum amplitudes. The minimum amplitude of a 50% AM tone is thus one-third the maximum amplitude. As the depth of modulation is decreased, the spectral energy at the carrier frequency grows and the energy at the sidebands decreases.

The amplitude of the steady-state response recorded at rates near 40 or 80 Hz increases as the depth of modulation increases. At both modulation frequencies, the amplitude saturates as the depth of modulation reaches about 50% (Picton et al, 1987c;

Lins et al, 1995). Part of this might be related to changes in the overall intensity of the stimulus as the modulation increases. If the rms intensity rather than the maximum amplitude is maintained constant (Viemeister, 1979; John and Picton, 2000b), the response is larger at 100% than at 50% (John et al, 2001b; Dimitrijevic et al, 2001). The phase is not significantly affected by changes in the modulation depth above about 25%.

The modulation may also follow an exponential sine function (John et al, 2002a). In order to use exponential envelopes for the amplitude modulation envelope, the formula for the amplitude modulation in the square brackets of the preceding formula becomes:

$$[1 + 2m_a(\{[1 + \sin(2\pi f_{am} t i)]/2\}^N - 0.5)]$$

where $N$ is the required exponent. An example of a tone modulated with an exponential sine function is given in Figure 8 (fourth line). A similar stimulus may be created by combining two AM tones with equal modulation frequencies and with carrier frequencies separated by twice the modulation frequency (Stürzebecher et al, 2001). When tones were amplitude modulated
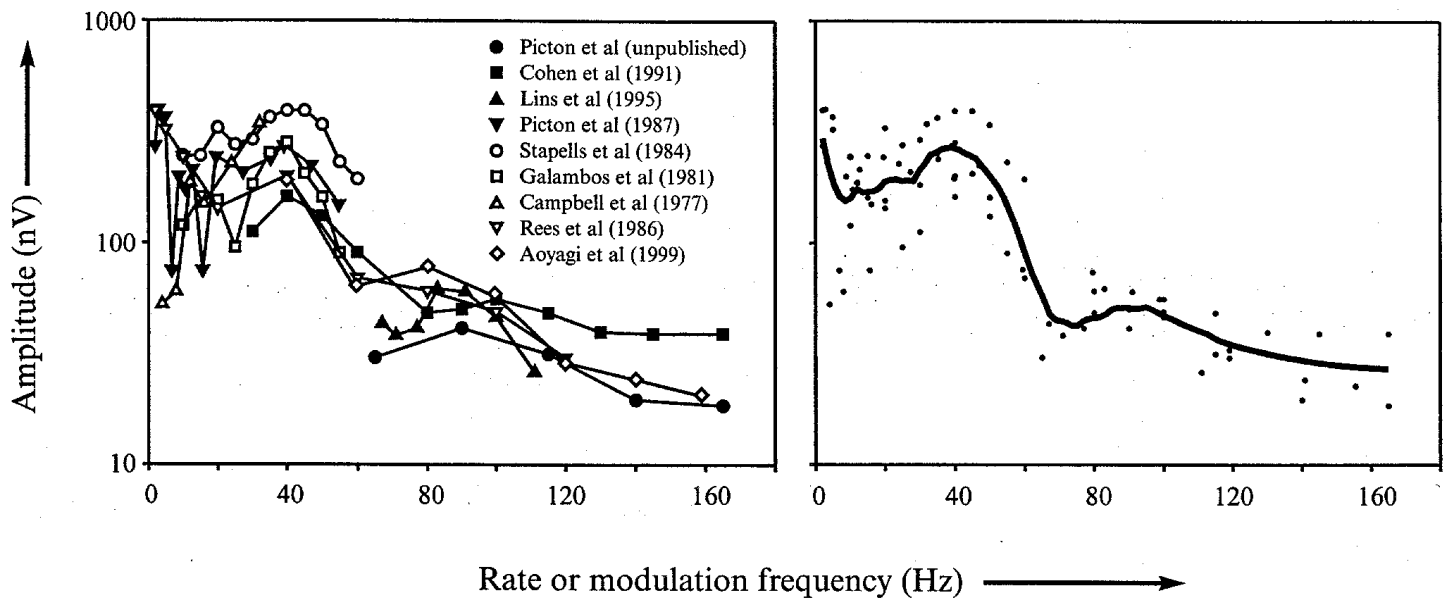
**Figure 10.** Effects of stimulus rate. The left graph shows data from nine different studies wherein the rate of stimulation or the modulation frequency was manipulated. The stimuli were all AM tones, except for the Galambos study, which used clicks, and the Stapells study, which used tone-bursts. Stimulus intensity varied between 55 and 70 dB HL. The data were made comparable using the following procedures. Peak-to-peak amplitudes have been halved to make them equivalent to baseline-to-peak measurements. The amplitudes of the responses to binaural stimuli have also been halved. The binaural effect varies with different stimulus rates, with binaural responses being about 200% of the monaural responses at rapid rates (Lins et al, 1995) and 150% at slow rates (Rees et al, 1986). Responses measured by, visual analysis in the time domain are usually larger than measured by computer analysis in the frequency domain, because the peaks in the time domain are enhanced by superimposed activity at other frequencies. Responses measured in the time domain were therefore multiplied by 0.75 to make them equivalent to those recorded in the frequency domain. Most of the results at frequencies less than 60 Hz are for waking subjects. If the subjects were asleep, these amplitudes would probably be reduced by a factor of 2 or 3 (Cohen et al, 1991). The actual stimuli in each of the experiments were: Picton et al (unpublished), AM tones at 60 dB SPL (measurements combined over multiple carrier frequencies); Cohen et al (1991), 55 dB HL tones, AM, binaural; Lins et al (1995), 1000-Hz AM tone at 60 dB SPL; Picton et al (1987c), 70 dB nHL AM 1000-Hz tones; Campbell et al (1977), 500-Hz AM tone; Stapells et al (1984), 60 dB nHL tone-bursts of 500 Hz; Galambos et al (1981), clicks of unspecified intensity; Rees et al (1986), 60 dB SL AM 1000-Hz tones; Aoyagi et al (1999), 50 dB HL AM 1000-Hz tones. The right graph plots the general trend of the various data sets plotted above. Each point on the smooth curve was fitted using a window of 22 points, with the points weighted using a bisquare function of the distance from the point being smoothed, and with the lines between the smoothed points interpolated using a second-degree polynomial.
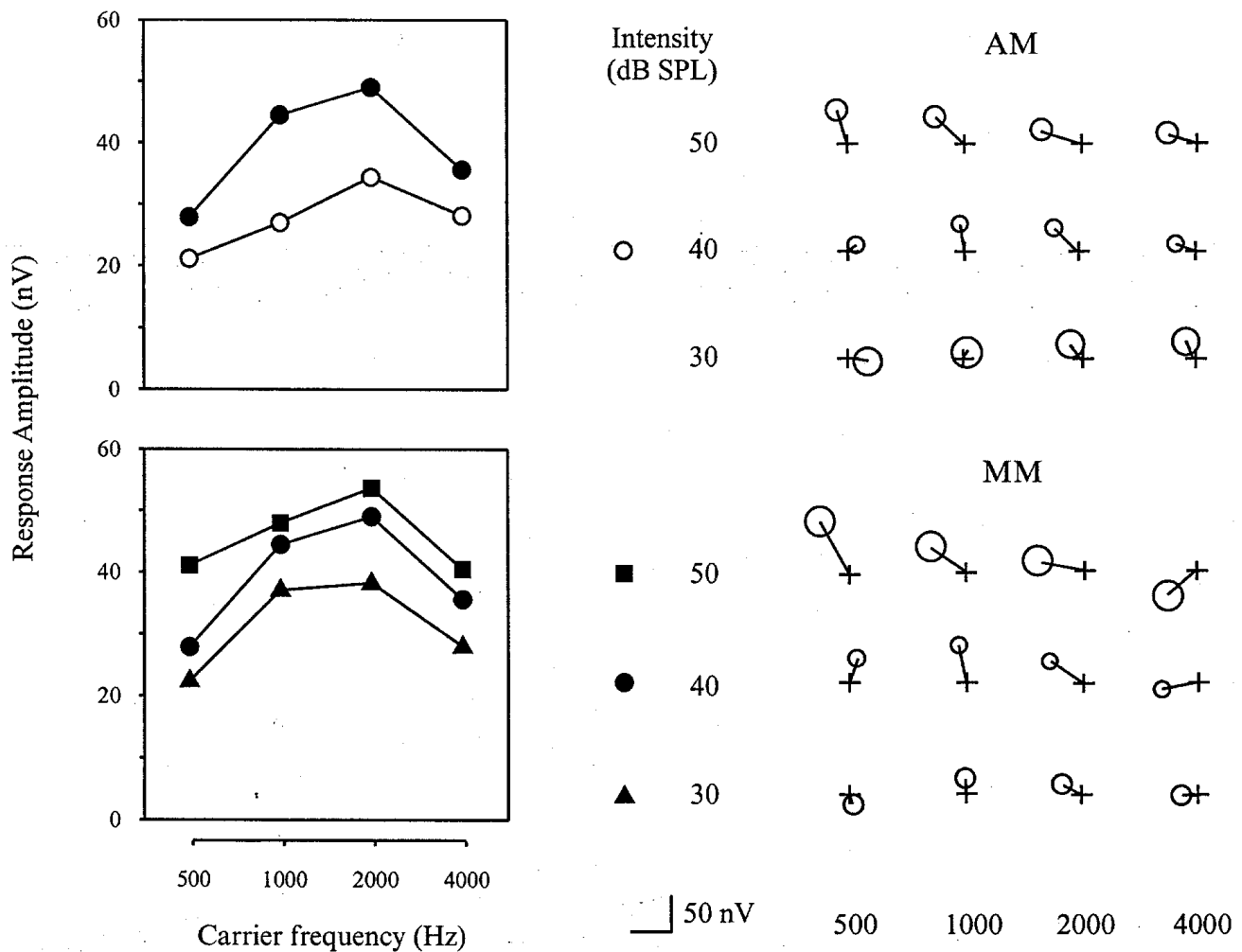
with exponential envelopes based on $\sin^N$, the amplitude and latency of the 80–100-Hz steady-state response increased significantly with increasing $N$. The increases were larger for lower and higher carrier frequencies, and smaller for carrier frequencies in the middle range of 1500–2000 Hz. These stimuli may therefore make the responses easier to recognize in noise. The change in the acoustic frequency specificity of the tone caused by the broader span of spectral energy is probably not clinically significant, given the frequency specificity of the response (see 'Frequency specificity', below).

*Intensity*

As the intensity of the stimulus increases, the amplitude of the response increases and the latency decreases. When brief tones of 400 or 500 Hz are presented at rates of 40 Hz, the amplitude (baseline-to-peak) of the response increases with increasing intensity, with a slope of approximately 5–10 nV/dB, and the latency decreases with a slope of approximately 100–200 μs/dB (Galambos et al, 1981; Stapells et al, 1984). The results are similar for tones of 1000 Hz, but the changes are much smaller for tones of higher frequencies (Rodriguez et al, 1986). The

responses grow regularly up to intensities of 90 dB HL. Similar changes are noted if the stimulus is the amplitude modulation of a continuous tone rather than separate tones, although the changes with intensity in both amplitude (3–9 nV/dB) and latency (80–100 μs/dB) are less (Picton et al, 1987c; Parker and O'Dwyer, 1998). At modulation rates near 80–100 Hz, the response is smaller and the amplitude change with intensity correspondingly less, approximately 2 nV/dB for intensities below 70 dB SPL (Lins et al, 1995). At higher intensities, the response shows a greater change with increasing intensity, and the harmonics in the response become more prominent (Lins et al, 1995). The latency change is 60–80 μs/dB, and is similar across carrier frequencies (John and Picton, 2000a). Figure 11 shows the effect of intensity on the 80–100-Hz responses.

The effects of intensity are mediated by multiple physiologic mechanisms. In the cochlea, the traveling wave will cause more intense stimuli to activate more cells, and there will be a basal movement of the location of these cells with increasing intensity. After the hair cells, synaptic delays will be affected by intensity. The increased effect of intensity on latency as the responses are generated at later parts of the auditory system is probably

Response Amplitude (nV)

60

40

20

0

60

40

20

0

500   1000   2000   4000

Carrier frequency (Hz)

Intensity (dB SPL)

AM

50

40

30

MM

50

40

30

50 nV

500   1000   2000   4000

**Figure 11.** Effects of carrier frequency and intensity. The responses plotted in polar format on the right side of this figure show the responses of one typical subject to AM stimuli and to mixed modulation (MM) stimuli. Mixed modulation combined 100% AM with 25% FM, with the relative phases between the modulations set so that the maximum amplitude and maximum frequency coincided. The four tones were presented simultaneously to the right ear at three different intensities. The response decreases in amplitude (the length of the line leading from the cross-hairs) as the intensity decreases. The onset phase (measured counter-clockwise from the $x$-axis) in general decreases with decreasing intensity. This is equivalent to the phase delay increasing. The circles represent the $p<0.05$ confidence limits of the response; if the circle does not include the origin (center of the cross-hairs), the response is significantly different from the background EEG. The graphs on the left side of this figure represent the mean data across eight subjects. The upper graph shows the difference between the AM responses (open circles) and the MM responses (filled circles) at 40 dB SPL. The lower graph shows the effects of intensity and carrier frequency on the MM responses. These data are derived from John et al (2001b).

caused by the increasing number of synapses in the pathway. The delay before the initiation of the next impulse at each synapse will vary with the intensity of the stimulus (which determines the number of presynaptic inputs and their rapidity). The greater amplitude changes with intensity at the slower modulation frequencies probably result from a combination of the increased number and synchronicity of the responding cells (in cortex as opposed to brainstem) and the shorter distance between the cells and the recording electrodes.

*Carrier frequency*

The effects of carrier frequency are quite different for stimuli modulated at rates near 40 Hz and near 80 Hz. The 40-Hz response significantly decreases in amplitude with increasing carrier frequency (Galambos et al, 1981; Rodriguez et al, 1986; Picton et al, 1987c). This phenomenon might be related to the

way in which the brainstem and cortical responses combine to give the scalp-recorded waveforms, since the 500-Hz brainstem response has a later and broader waveform that might more readily combine with the cortical response. Another possible explanation is that the carrier stimuli with lower frequencies have a broader activation pattern in the cochlea. More neurons are therefore activated, and this persists through the auditory system.

For the 80–100-Hz responses, the amplitude is larger for the mid-frequencies (1000–2000 Hz) than for either higher or lower frequencies (John et al, 2001b). Some of this effect at 80 Hz might be due to the stimuli being presented at the same SPL, whereas normal hearing thresholds are higher at the lower frequencies. However, the effect persists when stimuli are presented using constant HL rather than constant SPL (Dimitrijevic et al, 2002). The decreasing breadth of the activation pattern

might determine the fall-off at high frequencies. At low frequencies, the breadth of the activation pattern may have two effects. More neurons may be activated at the lower carrier frequencies, but these may respond with sufficiently different phases that the overall response might be attenuated.

The effect of modulation rate on the amplitude of the steady-state responses may vary with the carrier frequency and with the age of the subject. Cohen et al (1991) reported that the modulation frequency at which the auditory steady-state response was most efficiently recorded varied with the carrier frequency. However, these effects are not large in adults, and John and Picton (2000a) found no significant interactions between carrier frequencies (500–6000 Hz) and modulation frequencies (78–96 Hz). Rickards et al (1994) found that the response amplitude in neonates was larger at lower modulation frequencies for lower carrier frequencies, with optimal values of 72, 85 and 97 Hz for 500, 1500 and 4000 Hz, respectively. We have generally used lower modulation frequencies for lower carrier frequencies (Dimitrijevic et al, 2002), but this has been more related to the ease of identifying multiple responses in the spectrum than to any difference in amplitude across carrier frequencies.

*Frequency modulation and mixed modulation*
Frequency modulation involves changing the frequency rather than the amplitude of the carrier. Mixed modulation involves the simultaneous modulation of both the amplitude and the frequency of the stimulus. The basic formulae used to construct digitized versions of these stimuli are:

$$\varphi(i)=[m_f f_c/(2f_m)]\ \sin(2\pi f_m ti+\theta)$$

$$s(i)=a[1+m_a \sin(2\pi f_m ti)]\ \sin[2\pi f_c ti+\varphi(i)]$$

where $f_m$ is the frequency of modulation (of both amplitude and frequency), $f_c$ the frequency of the carrier, $m_f$ the amount of frequency modulation (from 0.0 to 1.0), $m_a$ the amount of amplitude modulation, $a$ the mean peak-to-baseline amplitude, $t$ the time per address in the stimulus buffer, $i$ a counter that goes from 1 to the size of the buffer, and $\varphi$ the term that adjusts the frequency of the carrier by changing its phase.

In these equations, the amount of frequency modulation is measured as the difference between the maximum and minimum frequencies divided by the carrier frequency. This is not the conventional way of measuring this parameter, which is usually expressed as $\beta$, the deviation from the mean carrier frequency divided by the modulation frequency. Percentage modulation expressed relative to the carrier frequency is easier to understand in audiometric terms, but, unfortunately, there is some ambiguity in this measurement. In the above equations, the percentage modulation is the full range of frequency change divided by the carrier frequency. If the frequency of a 1000-Hz carrier is modulated by 25%, the frequency changes between limits of 875 and 1125 Hz. This definition was used by Cohen et al (1991), and also Rickards et al (1994), and in all work from our laboratory (e.g. John and Picton, 2000b; John et al, 2001b; Dimitrijevic et al, 2002). Other work deriving from the original Cohen research (e.g. Rance et al, 1995; Rance and Briggs, 2002; Cone-Wesson et al, 2002a) has defined the percentage frequency modulation as one-half of this value, i.e. as the ratio of the deviation from the mean to the mean frequency.
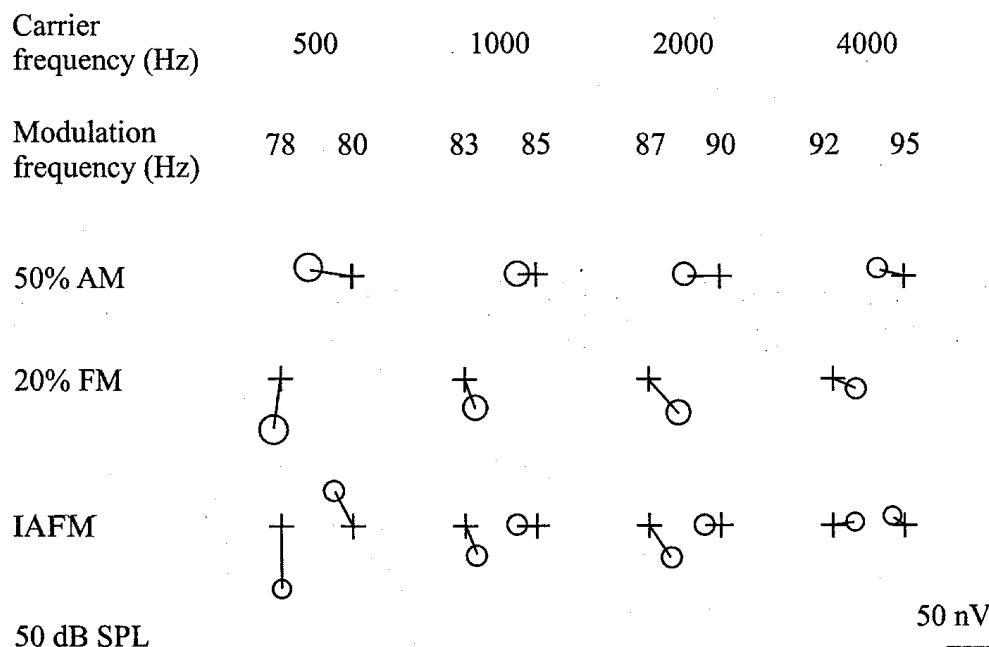
The relative phase between the amplitude and frequency modulations is given by $\theta$. This is arbitrary, but the literature generally considers the phase difference as zero when the stimulus reaches maximum amplitude and frequency at the same time. Such would occur if $\theta$ were set to –90° ($\pi/2$ radians) in the above formula. Another approach would be to change the above equations so that the trigonometric terms in the second line were cos instead of sin.

Frequency modulation can evoke auditory steady-state responses at both the fundamental frequency and at the second harmonic. The second harmonic response is a response to the rate of change of frequency, whereas the fundamental response may be the sum of separate responses in different frequency regions. At low frequencies of modulation (e.g. below 5 Hz), frequency modulation evokes large responses at the second harmonic, but at higher modulation frequencies, the response at the fundamental frequency is much larger than at the second harmonic (Picton et al, 1987b,c; Maiste and Picton, 1989).

Like the response to amplitude modulation, the response to frequency modulation is affected by the depth of modulation, the carrier frequency, and the stimulus intensity. At both 40 Hz and 80 Hz, the response to frequency modulation increases in amplitude with increasing depth of modulation but does not change its phase (Picton et al, 1987b,c; John et al, 2001b). The 40-Hz FM response is larger for lower carrier frequencies and larger for higher intensities (Picton et al, 1987b,c), but the effects of carrier frequency and intensity on the response to frequency modulation at 80–100 Hz have not been studied. For both the 40-Hz and the 80-Hz responses, the phase delay and apparent latency of the FM response is shorter than that of the AM response (Picton et al, 1987c; John et al, 2001b). This latency difference might be caused by interactions between the location on the basilar membrane where the modulation most effectively activates the neurons, and the different filter delays in the hair cell processing (John et al, 2001b).

The spectrum of a mixed-modulation stimulus varies with the relative phase between the two modulations. When the maximum amplitude and frequency occur at the same time, the spectrum has its maximum at a frequency above the carrier frequency (Figure 8). The effects of mixed modulation have been fully studied only for the responses at rapid modulation frequencies. Since the response to modulating a tone's frequency is relatively independent of the response to modulating its amplitude, the combined response to mixed modulation is a little less than the sum of the two responses (John et al, 2001b). The largest response occurs when the responses are in phase. This generally occurs when the maximum frequency occurs a little later than the maximum amplitude, although the optimum phase relationship between the two types of modulation varies with the carrier frequency.

Modulating a tone using different modulation frequencies for amplitude and frequency gives 'independent amplitude and frequency modulation' (IAFM). This allows two responses to be obtained for each carrier frequency. The responses are slightly smaller than when the modulations only involve amplitude or frequency, but still readily recognizable (Figure 12). They may provide an index of how well overlapping changes in frequency and amplitude are discriminated by the human nervous system (Dimitrijevic et al, 2001).

| Carrier frequency (Hz) | 500 | 1000 | 2000 | 4000 |

| Modulation frequency (Hz) | 78  80 | 83  85 | 87  90 | 92  95 |

50% AM

20% FM

IAFM

50 dB SPL

50 nV

**Figure 12.** Independent amplitude and frequency modulation (IAFM). These data are the responses of a typical subject to AM, FM and IAFM stimuli presented at 50 dB SPL. The upper two sets of data show the responses to 50% AM stimuli and to 20% FM stimuli. The phases of the stimuli were such that the maximum frequency occurred one-quarter of a cycle before the maximum amplitude. The responses are shown in polar format, with the onset phase of the response indicated by the angle of the response vector relative to the x-axis. If one were to adjust the timing so that maximum frequency and amplitude occurred at the same relative time, one would adjust the phase of the FM response by subtracting 90–180° (quarter-cycle to half-cycle clockwise). The lower set of data shows the responses when each carrier frequency was independently modulated in both amplitude and frequency. All responses remain significant, with little if any attenuation.

FM stimuli may also be used in more complex ways. Green et al (1979) described a stimulus that was frequency-modulated at a rate of 10 Hz, with a depth of modulation that was itself modulated at 4 Hz. This stimulus evoked clear responses at 4 Hz. Children with receptive language impairment show smaller responses than normal children (Stefanatos et al, 1989). Another type of stimulus involves the use of a regularly repeated pulse of frequency modulation (Stefanatos et al, 1997).

*Binaural processing*
The effects of binaural stimulation on the steady-state responses are quite different at different modulation frequencies. At rates of 40 Hz, the response to the monaural stimulus is only slightly smaller than the response to the same stimulus presented binaurally (Picton et al, 1985). This suggests that many of the neurons that generate the response are essentially binaural, responding equally well to stimuli in either ear. Presenting stimuli to both ears would then lead to physiologic occlusion, since the neurons cannot respond further than to the monaural stimulus. A less likely possibility is that different neurons respond to stimuli on the right and on the left and inhibit each other during binaural stimulation. The 80-Hz steady-state responses to binaural stimuli are larger than the responses to monaural stimuli, and the binaural response is approximately equal to the sum of the two monaural responses (Lins et al, 1995). This indicates that the neurons generating these responses are essentially monaural, and that the neurons responding to one ear interact little with those responding to the other ear. However, small interaction effects might be detectable if prolonged averaging is used.

Many binaural interactions other than the simple monaural versus binaural effects occur, but these have not yet been fully studied with the steady-state responses. Hink et al (1980) demonstrated binaural beats in the frequency-following response, but did not look at the envelope response to these beats. Wong and Stapells (2001) found no evidence for auditory steady-state responses to binaural masking level differences at rates of 80 Hz, but did find responses at 7 or 13 Hz. This work needs to be followed up.

*Multiple simultaneous stimuli*
Regan and Heron (1969) demonstrated that it was possible to record separate steady-state responses to each of four simultaneous visual stimuli. Each stimulus was presented to one quadrant of the visual field, and modulated at a rate different from the others. Responses to each stimulus could then be recognized by using four different Fourier analyzers, each locked to one of the stimulus frequencies. A similar approach can be used with the FFT, where the responses are detected in the specific frequency bins for each of the multiple stimuli. The technique basically involves three steps: presenting multiple stimuli with each having its own specific modulation frequency; accurately locking the analysis to each of the modulation frequencies; and using an analysis procedure with a frequency tuning that specifically measures each of the responses without contamination from the others. Regan and Regan (1988) were the first to record multiple auditory steady-state responses using pairs of AM tones with different modulating frequencies.

This approach can speed up the testing procedure by obtaining multiple responses in the time that would be necessary to record a single response. If the stimuli do not physically interact, if the responses are not affected by each other, and if the noise levels are unchanged by multiple as opposed to single stimuli, then the time required to record the responses is reduced by the number of stimuli presented simultaneously. Even if there is some reduction in the amplitude of the responses, it will still be more efficient to record the responses simultaneously if the amplitude reduction is less than $1\sqrt{N}$, where $N$ is the number of simultaneously recorded responses (John et al, 1998). Thus, if two stimuli are presented simultaneously and if the responses are greater than 70% of the amplitude when they are each presented singly, presenting the two stimuli together is more efficient than presenting them separately.

Lins and Picton (1995) first described the possibility of using multiple auditory steady-state responses to assess hearing at different frequencies and in different ears. They recorded steady-state responses using modulation frequencies between 80 and 105 Hz, and showed that, when the stimuli differed in ear and/or carrier frequency, there was no significant decrease in the amplitude of the responses for up to eight simultaneous stimuli (four in each ear). John et al (1998) found three limitations to this approach. First, there are significant interactions between carrier frequencies when they are separated by less than one half-octave. Second, the responses are less resilient to the presence of other stimuli when the intensities are increased to 75 dB SPL. Third, the responses recorded using modulation frequencies of 35–55 Hz showed much greater interactions than those recorded at higher modulation frequencies. These first two limitations are probably due to overlap of the activation patterns on the basilar membrane, which are broader at higher intensities. The third limitation indicates that the interaction between adjacent frequency regions is greater at higher levels of the auditory system such as the cortex, which contribute more to the scalp-recorded responses at lower modulation rates. At lower modulation frequencies (less than 55 Hz), the multiple stimulus technique was not significantly better than recording responses to each stimulus separately.

Dolphin (1996, 1997) has studied responses to multiple stimuli in dolphins and in gerbils. Instead of using sinusoidally amplitude-modulated tones, he created beats using pairs of stimuli. He found that using multiple two-tone pairs generally reduced the amplitude of the responses compared to the two-tone control condition. Each additional pair decreased the response by 1–3 dB (11–30% change in amplitude). This still makes it possible to record multiple responses more efficiently than single responses. Only when the reduction is more than 30% does it become less efficient. Furthermore, the envelope frequencies used in this experiment varied over a wide range from 38 to 142 Hz, and the maximum reductions with multiple stimulation were for the responses at the lower modulation frequencies (38 and 55 Hz). These findings are therefore in agreement with those of John et al (1998), who reported that, in humans, the multiple stimulus technique was much less efficient for the 40-Hz response than for the 80-Hz response. However, recent studies in our laboratory have shown that when multiple stimuli are presented simultaneously, the responses at the lower carrier frequencies may be slightly attenuated (John et al, 2002b). These effects are opposite to those that might be expected from masking, and may be related to phenomena such as suppression. In subjects with normal hearing, the response attenuation is small and does not override the time savings inherent in using multiple rather than single stimuli.

Picton et al (1998) demonstrated that, in aided individuals with impaired hearing, a steady-state response to a high-frequency stimulus could sometimes be recognized when that stimulus was presented alone, but not when it was presented concurrently with other stimuli having lower carrier frequencies. They attributed this finding to masking by the lower-frequency tones. Dimitrijevic et al (2002) found similar findings in some unaided patients. This problem probably occurs when auditory neurons have abnormally broad tuning curves. These masking effects will elevate the thresholds obtained using the multiple stimulus technique. However, the multiple stimuli may more closely approximate the sounds that occur outside of the clinic in the real world, where isolated pure tones are rare and most stimuli contain many different frequencies. The multiple stimulus assessment may therefore provide 'realistic' as opposed to 'optimistic' thresholds. The problem may only occur in some subjects, since Herdman and Stapells (2002) were able to obtain accurate thresholds in their patients with steeply sloping losses and did not find any difference between thresholds estimated using single and multiple stimuli.

As well as decreasing the amplitude of the responses, interactions between simultaneously presented stimuli may also enhance the responses. Dolphin and Mountain (1993) and Dolphin et al (1994) found that the responses in gerbils to tones with higher frequencies were slightly enhanced by the presence of the low-frequency tones. A similar small effect has also been found in human studies (John et al, 1998, 2002b).

Recording responses to multiple simultaneous stimuli can also demonstrate the non-linear distortions that occur during the processing of the stimuli. Two simultaneous sinusoidal stimuli with frequencies $f_1$ and $f_2$ will evoke from a non-linear system a series of distortion products at frequencies $nf_1 \pm mf_2$, where $n$ and $m$ are integers (Regan and Regan, 1988, 1993; Regan, 1994a,b). Distortion products resulting from processing in the cochlear hair cells can be recorded acoustically as otoacoustic emissions, and Chertoff and co-workers (Chertoff and Hecox, 1990; Chertoff et al, 1992) demonstrated that such frequencies can be measured electrically as well as acoustically. Using combinations of two tones $f_1$ and $f_2$, they demonstrated responses at both the $f_1$–$f_2$ difference tone (or beat frequency) and the $2f_1$–$f_2$ cubic difference tone. The pattern of the electrical distortion products can be used to characterize the types of non-linearity that occur in the processing of the stimuli. Regan and Regan (1993) showed that the pattern of responses to two AM stimuli could be predicted by the hair cell transducer function, which is essentially a compressive rectifier. Lins and Picton (1995) used this approach to evaluate a model of the 80-Hz response consisting of a compressive rectification (cochlea) and a filter that passes different ranges of modulation frequency (brainstem). These approaches need to be evaluated more extensively.

*Sweeping through stimulus parameters*
Regan (1973, 1989) demonstrated how to record the effects of a particular stimulus parameter on the steady-state responses by sweeping through different values of that parameter rather than

by making multiple separate recordings at different parameter settings. The sweep technique is relatively simple when using a Fourier analyzer, since the output of the analyzer can react quickly enough to follow the signal that is sweeping the stimulus parameter. One can also use short-time FFT analyses and repeat these through the sweep. Furthermore, one can repeat the sweep many times and average the resultant graphs of response versus stimulus parameter. If the general shape of the graph is known beforehand, the averaged graph can be smoothed to fit this shape. If some characteristic of the graph (such as an inflection at threshold or a maximum value) is more important than the individual points in the graph, the sweep technique may be much more efficient than multiple individual measurements. One difficulty with the technique is that the frequency resolution of the analysis will depend upon the low-pass filter used to track the changes in the amplitude of the response and that this might lead to increased noise in the recordings. Sweep techniques have been used extensively to estimate thresholds for visual discriminations (Norcia et al, 1989).

Sweep techniques have been used to graph the effect of intensity (Linden et al, 1985; Rodriguez et al, 1986) and stimulus rate (Linden et al, 1985; Sapsford et al, 1996) on the auditory steady-state responses at rates near 40 Hz. Rodriguez et al (1986) found that estimating thresholds was more effectively performed by separate evaluations at different intensities than by graphing the effect of intensity over a whole range, since testing separately at different intensities could more quickly home in on the threshold levels. Dobie and Wilson (1995b) found that an 'adaptive' selection of intensities with increased averaging at lower intensities was more effective than performing a full intensity series or estimating thresholds using regression. We have recently begun to use sweep techniques with the auditory steady-state responses evoked by more rapidly presented stimuli. Figure 13 shows some of our initial data. The right side of the figure shows how threshold might be estimated by regression analysis of the amplitude data significantly different from the residual background noise. The extrapolation is based on the assumption that the amplitude near threshold is linearly related to the intensity (measured in dB). Norcia et al (1989) review how to handle data that show non-linear relationships with intensity. Threshold estimation might be improved by using the phase (both its magnitude and its variance) in addition to amplitude. The sweep technique can be used with multiple as well as single stimuli. The data in Figure 13 were obtained with concurrent stimuli in the left (at 85 Hz) and right (at 95 Hz) ears.

## Intracerebral sources

*Physiologic introduction*
A full evaluation of how the auditory nervous system responds to sustained modulations of sound is not within the purview of this paper. However, a brief summary of what occurs during the processing of these sounds will provide a basis for discussing the intracerebral sources of the scalp-recorded activity. For simplicity, we shall focus on the sinusoidal amplitude modulation of a pure tone at a low or moderate intensity. This stimulus contains no energy at the modulation frequency, and activates the basilar membrane in the region specific to the carrier frequency.
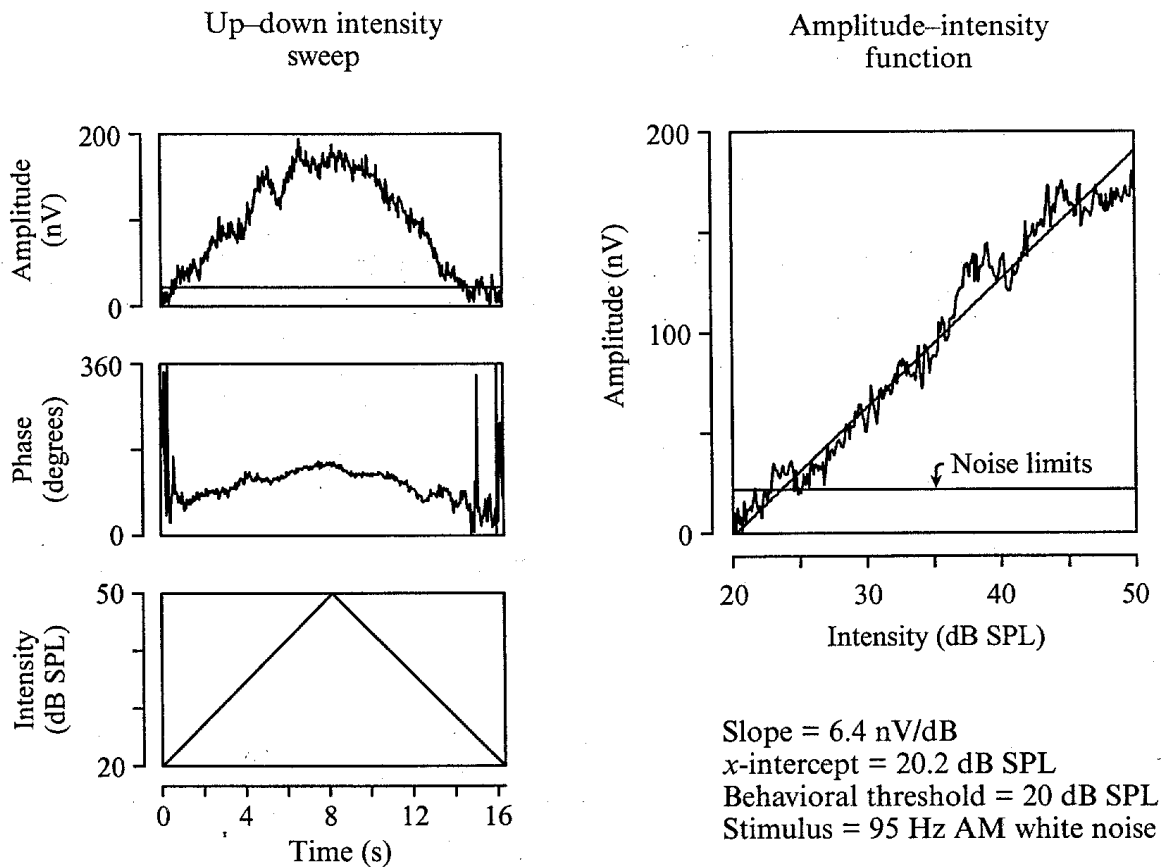
The hair cell transduction is not linear, but involves compression as intensity increases and an asymmetry of response to hair movements in different directions. At the synapse, the asymmetry becomes a full rectification, since the afferent fiber is only activated by hair cell depolarization. These non-linearities (compression and rectification) cause the neural response to carry energy at the modulation frequency as well as at the carrier frequency. Because neural firing rates are limited, the carrier frequency can only be followed at lower frequencies. A frequency-domain analysis of the post-stimulus time histogram of the response of an auditory nerve fiber to an AM or FM tone shows components at the modulation frequency (and harmonics) as well as at the carrier frequency and its sidebands (Khanna and Teich, 1989a,b). Envelope following therefore occurs as early as the primary nerve fibers.

The brainstem auditory nuclei maintain this information. The inferior colliculus is particularly responsive to the modulation frequencies of stimuli, with collicular neurons showing transfer functions that are tuned to the modulation frequency (Langner and Schreiner, 1988; Krishna and Semple, 2000). 'Best modulation frequencies' occur at frequencies up to several hundred hertz. Fastl (1990) pointed out that these transfer functions may be related to the psychophysical perception of the 'roughness' of a modulated pure tone. Although FM sensitivity in the auditory nerve is probably a mapping of how the activation pattern moves back and forth on the basilar membrane, i.e. mediated by multiple AM-sensitive neurons, neurons responding specifically to frequency modulation occur in the brainstem. Neurons in the inferior colliculus can respond specifically to modulations of either frequency or amplitude (Poon and Chiu, 2000). The neurons of the auditory cortex show a multiplicity of response patterns (review: Eggermont, 2001). They may respond to particular intensities as well as frequencies, and they may respond to specific directions and speeds of modulation.

Envelopes provide important information. They help in the localization of a modulated sound source. Interaural phase delays and interaural intensity differences for the modulating frequency are as important as those characteristic of the carrier frequency. The envelope is also essential to the determination of auditory objects. Sound components that are modulated in the same way probably originate from the same object. Eggermont (2001) has recently suggested that the auditory system might process complex sounds using two main processes: one mapping the 'texture' of a sound (those aspects of the spectrum that are important for object identity), and the other synchronizing the analysis to the 'contours' of a sound (onsets, offsets, and slow modulations). Auditory steady-state responses provide a way to look at contour processing.

*Latency measurements*
The latency of transient auditory evoked potentials is important when determining their intracerebral sources. Although the auditory system has both ascending and descending connections, the synchronized potentials recorded at the scalp seem to be generated mainly in the ascending system. Therefore, the later the latency of a wave, the higher in the auditory system is its generator. Unfortunately, as reviewed in 'Phase and latency', above, the latency for steady-state responses can only be estimated from measurements such as the 'apparent latency' or 'group delay'.

## Up–down intensity sweep

Amplitude (nV)

200

0

Phase (degrees)

360

0

Intensity (dB SPL)

50

20

0  4  8  12  16

Time (s)

## Amplitude–intensity function

Amplitude (nV)

200

100

0

Noise limits

20  30  40  50

Intensity (dB SPL)

Slope = 6.4 nV/dB
x-intercept = 20.2 dB SPL
Behavioral threshold = 20 dB SPL
Stimulus = 95 Hz AM white noise

**Figure 13.** Auditory steady-state responses to a sweep of intensity. On the left of the figure are shown the amplitude and phase of the response to a 95-Hz AM white noise in a single subject. The intensity of the stimulus was increased from 20 to 50 dB SPL over the first half of the 16.384-s sweep, and then decreased back down to 20 dB SPL over the second half. The data were averaged over 64 sweeps (approximately 17 min). In the first and last seconds of the recording, the amplitude tends to be below the $p < 0.05$ confidence limits (determined by the background activity at adjacent frequencies) shown by the horizontal line. During this time, the phase is quite variable. In the middle portion of the recording, the amplitude increases and the onset phase increases (or phase delay decreases) with increasing intensity. On the right of the figure, the amplitude values from the upper left graph have been collapsed over the 20–50-dB intensity range by vector averaging. A regression line calculated on those amplitude values greater than the noise limits can be extrapolated down to an estimated threshold at the x-intercept of 20.2 dB SPL, very close to the subject's behavioral threshold of 20 dB SPL.

The apparent latency of the steady-state responses varies with the modulation frequency and the carrier frequency. Rickards and Clark (1984) reported that the apparent latency decreased regularly with increasing modulation frequency. This may be related to the fact that the later stations of the auditory system respond less well to more rapid stimuli. Table 1 shows the apparent latencies reported in the literature for the auditory steady-state response.

The latency estimates in Table 1 vary significantly. Some of the variability may be intrinsic to the measurement. Picton et al (1987c) reported a range of 32–50 ms across 10 subjects for the AM response, which is similar to the standard deviation of 6 ms reported by Stapells et al (1987) for brief tones at 500 Hz. This might result if multiple generators (each having different latencies) contribute to the response and if the predominant generator varies from one individual to the next. At frequencies greater than 80 Hz, the standard deviations of the latency estimates are smaller (of the order of 1 ms in John and Picton, 2000a), perhaps because there is one dominant generator in all subjects.

Many of the differences in Table 1 may relate to the different intensities of the stimuli. Intensity has several effects. A main effect is that a more intense stimulus causes synapses to initiate impulses more rapidly, and this effect may accumulate over a multi-synaptic pathway. Another effect is the basal shift in the location of the effective stimulus at higher intensity. John and Picton (2000a) found a mean decrease in latency of 2.4 ms across carriers for responses in the 90-Hz range, when the stimulus level was changed from 35 dB SPL to 75 dB SPL. This can be converted to a slope of 0.06 ms/dB, identical to the Lins et al (1995) value for the intensity effect at this response frequency range. If one considers the effects of intensity and intersubject variability, the only real discrepancy in Table 1 is the short Cohen et al data for modulation frequencies of greater than 90 Hz. This is probably because their calculations included higher modulation frequencies (often into the 170-Hz range).

The latency of the responses decreases consistently with increasing carrier frequency. This is probably related to the traveling wave in the cochlea. There is some suggestion in Table 1 that the delay from high to low frequencies may be less when the stimulus intensities are greater, and this may be related to some basal movement of the dominant cochlear locations for low frequencies at higher intensity. However, the traveling wave

**Table 1.** Estimated latencies of auditory steady-state responses at different stimulus rates

| Report | Stimulus rates (Hz) | Stimuli[a] | Intensity (dB HL)[b] | Latencies (ms) at different carrier frequencies (Hz) | | | |
|---|---|---|---|---|---|---|---|
| | | | | 500 | 1000 | 2000 | 4000 |
| **40-Hz range** | | | | | | | |
| Stapells et al (1984) | 35–55 | TB | 70 | 29 | 25 | 20 | 20 |
| Kuwada et al (1986) | 25–60 | AM | 80 | | 31 | | |
| Stapells et al (1987) | 29–54 | TB | 50 | 41 | | | |
| Picton et al (1987c) | 27–55 | AM | 70 | | 37 | | |
| | | FM | | | 38 | | |
| Cohen et al (1991) | 30–60 | MM | 55 | 33 | 25 | 29 | 29 |
| **90-Hz range** | | | | | | | |
| Cohen et al (1991) | Above 90[c] | MM | 55 | 13 | 13 | 9 | 9 |
| Lins et al (1995) | 67–111 | AM | 50 | | 19 | | |
| John and Picton (2000a) | 80–100 | AM | 50 | 22 | 19 | 17 | 16 |
| John et al (2001b) | 75–95 | AM | 40 | 31 | 25 | 15 | 13 |
| | | FM | | 28 | 22 | 12 | 18 |
| Purcell et al (2003)[d] | 80-92 | B | 30 | | 22 | 16 | 18 |
| **170-Hz range** | | | | | | | |
| Kuwada et al (1986) | 150–350 | AM | 80 | | 8 | | |
| John and Picton (2000a) | 150–190 | AM | 50 | 12 | 10 | 8 | 7 |
| Purcell et al (2003)[d] | 177–199 | B | 30 | | 18 | 11 | 10 |

All latencies (ms) estimated by calculating the apparent latency from the slope of the phase versus frequency function, except for the data in John and Picton (2000a), which were calculated using the preceding-cycles technique. Apparent latencies in this study were slightly slower (exact data were not available for tabulated carrier frequencies).

[a] Stimuli: tone-bursts (TB), amplitude modulation (AM) or mixed modulation (MM), beats from tone pairs (B).
[b] Intensity: estimated in dB HL from the reported SPL and SL measurements.
[c] The upper range varied with the carrier frequency from 115 at 500 Hz to 185 at 4000 Hz.
[d] The Purcell study used beats with frequencies near 900, 1800 and 3600 Hz.

effects are small, and changes in these are not always easy to demonstrate, given the intersubject variability.

Otoacoustic emission latency can also be estimated by using the apparent latency technique. Measurements of the change of emission phase with response frequency in the ear canal can be used to determine the round trip delays of the stimuli to the place of emission generation and back to the acoustic microphone. It might be helpful to use otoacoustic emission latency for estimating the peripheral part of the evoked potential delay. This might then allow the neural transmission time of the evoked potential to be separated out. Distortion-product otoacoustic emissions are elicited using long-duration tone pairs, and these stimuli can concurrently evoke a steady-state potential to the beat frequency of the tone pair (Purcell et al, 2003). Unfortunately, the exact forward latency for distortion-product otoacoustic emissions is difficult to determine. Since stimulus and response frequencies are different, and since multiple sources on the basilar membrane can be active, the forward latency is not simply half of the total (Knight and Kemp, 2000).

*Multichannel recordings*
Auditory steady-state responses are widely recorded over the scalp. Johnson et al (1988) presented the first scalp map of the 40-Hz auditory steady-state response, showing clear polarity inversions over both temporal lobes in some subjects. This mapping was based on the voltages of one particular peak in the time-domain response. When the data are in the frequency domain, the topography of the steady-state response can be displayed by using multiple polar plots. The arrows show the amplitude and phase of the response at different scalp locations, in much the same way as when meteorologists map the intensity and the direction of the wind. Figure 14 shows wind-vane scalp topographies for the steady-state responses to AM tones at 39 Hz. The largest responses are recorded from the frontocentral regions and from the posterior neck. The phases of these responses are different by approximately 180°. The best one-channel montage would thus be between a midfrontal or vertex electrode and the posterior neck. However, the neck electrode may be affected by muscle noise when the subject is awake, and the recording between the vertex and the inion (where the response is only a little smaller than at the posterior neck) may then provide a better signal-to-noise ratio (van der Reijden et al, 2001).

Contour-mapping the relative amplitudes across the scalp is difficult to understand, since the amplitude will be equivalent for responses with inverted polarity. Lehmann and Michel (1989, 1990) showed how to calculate a dominant-phase axis from the slope of the regression line of the imaginary values versus the real values (using the paired values at each electrode). Mapping can then be performed by projecting the responses onto this dominant-phase axis. This works well if the scalp responses show a dominant phase, i.e. are all aligned along one phase or its inverse. This would be true if the response was generated by a single dipole. Figure 15 shows how this procedure works with the 39-Hz data from Figure 14. If the scalp response is generated by more than one intracerebral generator, and if these
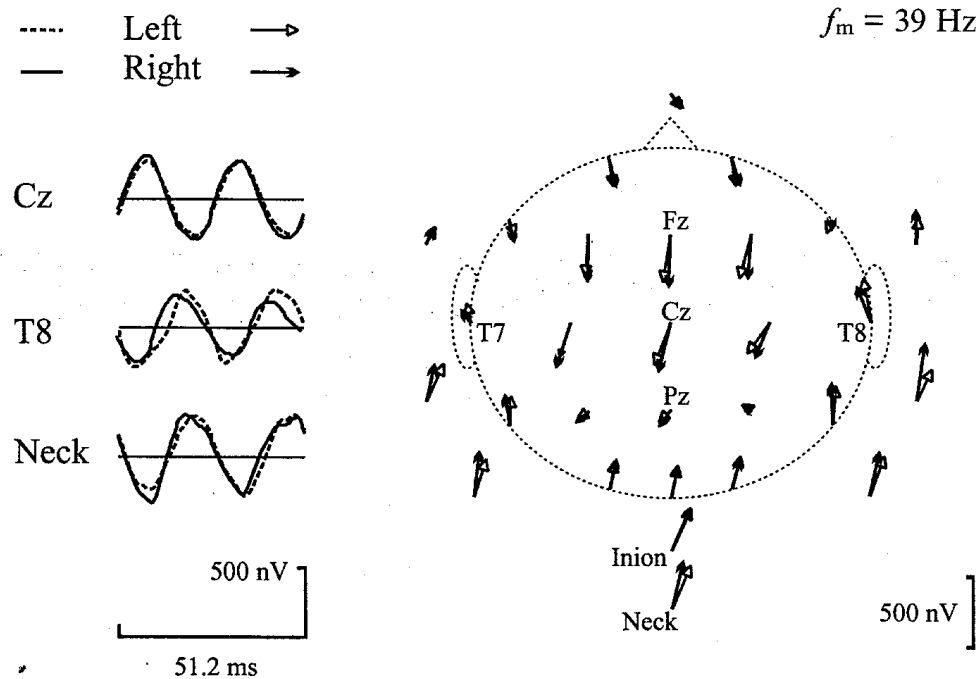
**Figure 14.** Scalp topography of the auditory steady-state responses. On the left of the figure are shown the scalp-recorded responses at three electrode locations (out of 47 total). The stimulus was a 1000-Hz tone amplitude-modulated at 39 Hz and presented to either the left or right ears at 70 dB SPL. The responses are averaged over 10 subjects, and shown over a two-cycle period. The subjects were awake throughout the recordings. The responses at the vertex (Cz) are approximately the inverse of the responses at the posterior neck, i.e. they differ in phase by about 180°. The responses at the right midtemporal location (T8) show a different phase from those at the vertex. In the right part of the figure are shown the polar plots of these responses and those at other scalp locations. These polar plots show the amplitudes and phases of the response at the different scalp locations. Conversion of the time-domain waveforms to the polar plots is illustrated in Figure 1 for responses to a stimulus with a slightly different rate, recorded from one subject rather than averaged over 10. The polar plots are arrayed over the scalp (with the beginning of the arrows at the electrode locations), and viewed from above with the nose at the top of the figure. The dotted-line ellipse represents the 'equator' of the cephalic sphere passing through electrode locations Fpz, T7, Oz, and T8. Electrode locations below the equator are plotted outside the ellipse. The responses are plotted such that the beginning of the arrows is at the electrode location. The length of the arrow is equivalent to the amplitude of the response, and the counter-clockwise angle of the arrow relative to the horizontal axis is equivalent to the cosine onset phase of the response. The phase delay would be measured in the clockwise direction. Because the stimulus was modulated using a sinusoidal envelope and the response phase is measured with respect to a cosine reference, the actual phase delay is 90° less than that shown (and the polar plots could all be rotated a quarter-cycle counter-clockwise). This figure shows data that are more extensively described in Herdman et al (2002a).

generators have different phase relations to the stimulus, the contour map will vary through the cycle of the response, and the dominant-phase projection will miss this variability (Herdman et al, 2002a).

*Source analysis*
The sources for the electrically recorded steady-state responses have not been extensively studied. Johnson et al (1988) performed a 21-channel mapping study and demonstrated clear polarity inversions of the response to 40-Hz brief tones over the midtemporal regions in some of their subjects. They suggested that the 40-Hz response might be generated in the auditory cortices and in thalamocortical circuits.

Intracerebral sources for the steady-state responses can be estimated using dipole source analysis, either directly in the frequency domain (Lütkenhöner, 1992), or by treating the real and imaginary components at the response frequency as independent values (Herdman et al, 2002a). Mauer and Döring (1999) reported that both brainstem and cortical (temporal

lobe) sources are active during the amplitude modulation following response (using modulation frequencies between 24 and 120 Hz). However, the intensity of cortical activity decreased with increasing modulation frequency, and the brainstem was the dominant source at modulation rates of greater than 50 Hz.

Herdman et al (2002a) have proposed similar intracerebral sources for the scalp-recorded auditory steady-state responses. A continuous 1000-Hz tone presented to the right or left ear at 70 dB SPL was sinusoidally amplitude-modulated to a depth of 100% at 39 or 88 Hz. Sample data from 47 electrodes on the head are shown in Figure 14. The main source model contained a midline brainstem generator with two components (one vertical and lateral) and cortical sources in the left and right supratemporal planes, each containing tangential and radial component (Figure 16). At 88 Hz, the greatest activity occurred in the brainstem, and subsequent cortical activity was minor. At 39 Hz, the initial brainstem component remained, and significant activity also occurred in the cortical sources, with the tangential activity being larger than the radial. These results

+400

0

−400

−400          0          +400 nV

φ

Voltage
0.1 µV
contours

CSD
0.069 µV/cm²
contours
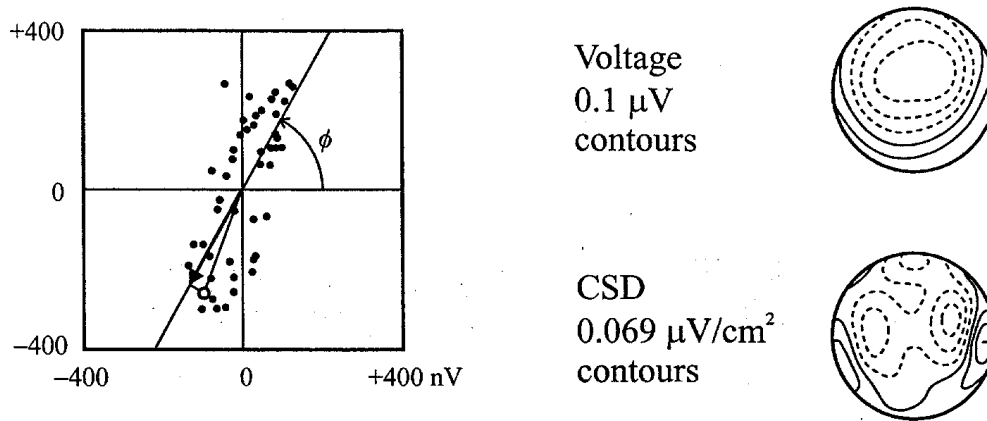
**Figure 15.** Contour-mapping of a multichannel steady-state response. The left part of this figure shows the two-dimensional data for the 39-Hz response to left ear stimulation in Figure 14. The open circle represents the data from the vertex electrode. The slope of the regression line shows the predominant phase (ø) of the response (61°). The data from each electrode are then projected onto the axis (illustrated for the vertex data). The graphs on the right then show the projected data mapped onto the scalp using contour-mapping. Maps are shown for both voltage and current-source density (CSD), which measures the slope of the surface voltage. Contours for the positive values are shown with continuous lines, and contours for the negative values are shown with dashed lines. The outer edge of the map extends beyond the equator to the level of the mastoids. The vertex response has a phase of −111° and therefore projects onto the axis with a negative value. The pattern of the CSD scalp topography suggests dipole sources in the left brainstem and the right temporal cortex.

suggest that the whole auditory nervous system is activated by modulated tones, with the cortex being more sensitive to lower modulation frequencies. Figure 16 presents some of the source activities both as temporal waveforms and as polar plots.

*Magnetoencephalographic (MEG) recordings*
MEG has some definite advantages over electrical recordings for determining sources. Since MEG recordings are unaffected by the impedance of the volume conductor in which the sources are located, they provide more precise localization of sources. MEG fields recorded at the scalp are mainly generated by current sources oriented tangentially to the surface, are unaffected by radially oriented sources, and are relatively insensitive to deep sources. These characteristics significantly simplify the analysis, since one does not have to disentangle the overlapping effects of many sources. In the auditory system, MEG recordings therefore indicate mainly what is happening in the auditory cortices located on the superior temporal plane. Other tangential generators, such as those in the walls of the superior temporal gyrus, may also contribute to the recordings.

The auditory steady-state responses evoked by stimulus rates near 40 Hz have been extensively studied using MEG. The MEG response to AM tones is affected by such basic parameters as modulation frequency, intensity, and carrier frequency (Roß et al, 2000), in a similar manner to the electric responses (e.g. Picton et al, 1987c). The magnetic response has its maximal amplitude for modulation frequencies near 40 Hz, increases with increasing intensity according to a linear function up to 80 dB SL, increases with increasing depth of modulation up to 80%, and decreases with increasing carrier frequency. The apparent latency of the response was 48 ms for 250-Hz tones. The similarity of these results suggests that the scalp-recorded electric response is, like the magnetic response, largely generated in the supratemporal plane.

Early studies of the 40-Hz steady-state responses suggested that their sources were similar to the sources of the transient

middle-latency responses (Mäkelä and Hari, 1987), and that the steady-state response could represent the overlapping of the transient responses (Hari et al, 1989). However, later studies found that the transient and steady-state sources could localize to different parts of the superior temporal plane (Ciulla et al, 1996). Furthermore, the tonotopy of the 40-Hz responses was organized with the sources for low-frequency tones located lateral to the sources for high-frequency tones, quite the inverse of the tonotopy of the transient middle-latency response (Pantev et al, 1996). Different components of the transient middle-latency response have different source locations (Yvert et al, 2001). These may show complicated interactions when they overlap to form the steady-state responses (Gutschalk et al, 1999).

There has not been much MEG study of the steady-state response at more rapid modulation frequencies. Roß et al (2000) reported a secondary response maximum at stimulus rates near 80 Hz, with an apparent latency of 26 ms for these responses. These responses were found in the superior temporal plane. Schoonhoven et al (2001) also found similar temporal sources for both 40-Hz and 80-Hz responses. They suggested that different regions of the auditory cortex might contribute at the different rates. Unfortunately, MEG is relatively insensitive to deep sources, and Schoonhoven et al did not consider additional deep sources.

*Nature of the auditory steady-state responses*
There are complex relationships between the steady-state responses and the spontaneous background rhythms of the EEG. The steady-state response may be in part a driving of the background activity by the stimulus and in part an independent response. In either case, many of the same neurons are probably involved in both the steady-state response and the background rhythms.

In their initial description, Galambos et al (1981) suggested that the auditory 40-Hz response to repeated clicks or tones
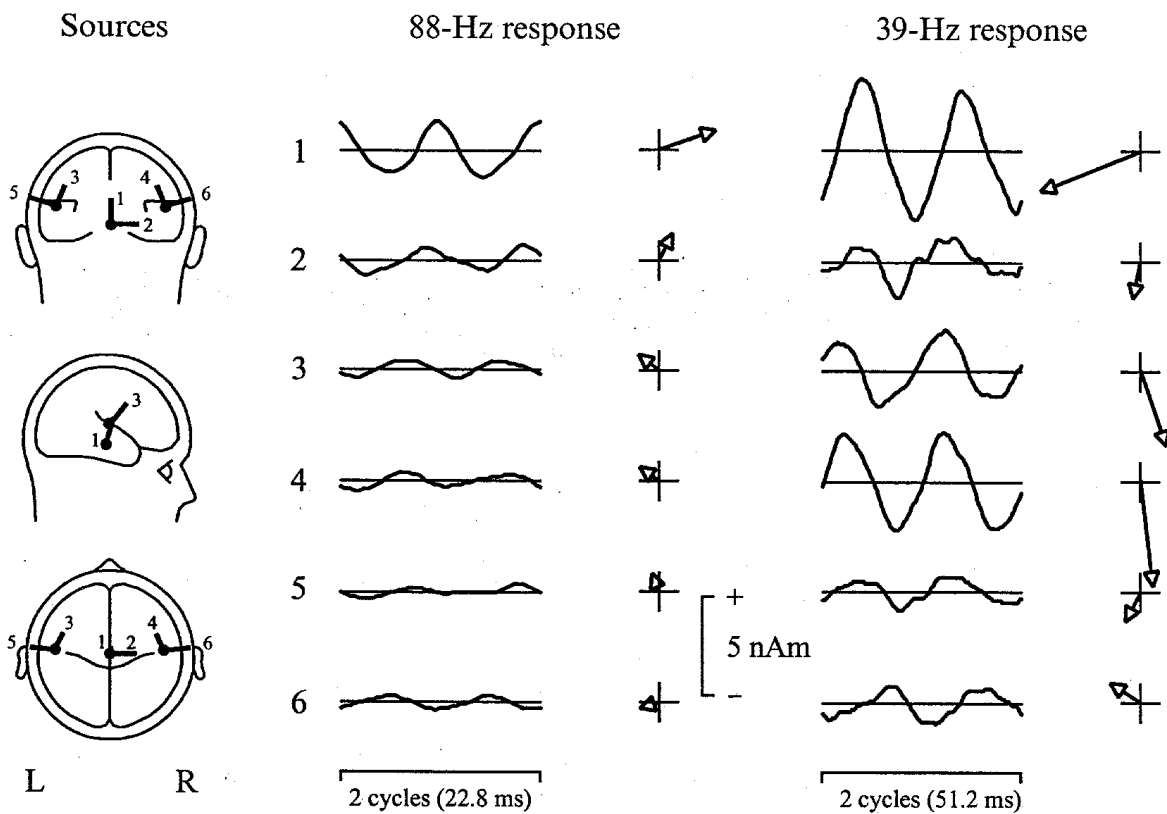
**Figure 16.** Sources of the auditory steady-state responses. This figure shows the source analysis of the data for left ear stimulation at 88 and 39 Hz. The surface recordings that led to the results for the 39-Hz responses are shown in Figures 14 and 15. The left of the figure shows the locations and orientations of the sources in the proposed solution. These are plotted using unit vectors plotted in three orthogonal planes. The three-dimensional orientation of sources 3 and 4 is upward and towards the midfrontal region of the scalp. The source activity for each response is shown in both the time domain (for a response collapsed over two cycles) and as a polar plot. The polar plots are plotted using double the scale of the temporal plots (i.e. the length of the vector is roughly equivalent to the peak-to-peak amplitude of the temporal waveforms). The polar plots are graphically similar to the source orientations, but show the amplitude and phase of the source waveform rather than the three-dimensional orientations of the sources. The brainstem (1 and 2) is the major source for the 88-Hz response. The tangential cortical sources (3 and 4) are also active, and these activations likely occur later than the activity in the vertically oriented brainstem source—the waveforms are shifted to the right. The main sources of the 39-Hz response are the brainstem (1) and the tangential cortical sources (3, 4). There is also some activation of the radial cortical sources (5, 6) occurring a little later than the activity in the tangential sources. This figure shows data that are more extensively described in Herdman et al (2002a).

represented the superimposition of the transient middle-latency responses to each of the stimuli. Since the waves of the transient response recurred at inter-peak intervals of about 25 ms, the superimposition was most effective at stimulus rates of 40 Hz. Several studies using electric (Galambos and Makeig, 1988; Plourde et al, 1991; Stapells et al, 1988) or magnetic (Hari et al, 1989; Gutschalk et al, 1999) recordings supported the idea that the steady-state response represents superimposed transient responses.

The prediction of the steady-state response from the superimposition of the transient response waveform is, however, not always accurate. Azzena and co-workers attempted to predict the steady-state response to trains of stimuli from the response to single stimuli. Although the prediction was close at rates of 40 Hz, significant differences occurred at higher or lower stimulus rates (Azzena et al, 1995; Santarelli et al, 1995; Conti et al, 1998; Santarelli and Conti, 1999). The recorded steady-state response is therefore, to some extent, independent of the transient response. One possibility is that it represents the

entrainment of a neuronal rhythm (Basar et al, 1987), instead of or in addition to the superposition of the responses. Another possibility is that the waves of the transient middle-latency response might show temporal refractory effects at higher rates (Conti et al, 1999). This might explain the tendency for the predicted response to be larger than the actual steady-state response at the higher rates. Several different mechanisms probably participate in the generation of the responses to brief trains of stimuli. Roß et al (2002) demonstrated that the onset and offset of a train evoked MEG gamma-band responses as well as a steady-state response. Furthermore, the development of the steady-state response at the onset of a modulated tone occurred over a much longer time (about 200 ms) than its resolution at the offset of the tone.

The nature of the 80-Hz steady-state response has not been extensively considered. Lins et al (1995) suggested that it might represent the superimposition of several auditory brainstem responses (predominantly wave V) evoked by each increase in the amplitude of the stimulus, or some independent response

evoked specifically by the modulation signal. Wave V of the auditory brainstem response has many of the same characteristics as the 80-Hz steady-state response, including the topography and the responsiveness to different tonal frequencies. Kiren et al (1994) showed in the cat that the envelope following response at all rates is markedly reduced by bilateral lesions to the inferior colliculus. However, the latency of the steady-state response is longer than one would expect for the brainstem generators. This may be related to multi-synaptic pathways in the brainstem, or the scalp-recorded response may be a combination of cortical and brainstem responses, as suggested by the source analyses.

Several other rhythmic responses can also be evoked by an auditory stimulus. The 'frequency-following response' is a brainstem potential that follows the acoustic pattern of the stimulus (Moushegian et al, 1973). This brainstem response is best recognized with carrier frequencies of 1000 Hz or less and at intensities 40 dB or more above threshold (Batra et al, 1986; Parker and Matsebula, 1998; Krishnan and Parkinson, 2000). Studies of this response to multiple tones or pseudo-random noise demonstrated that the envelope of the stimulus could evoke larger brainstem responses than the carrier (Hall, 1979; Dobie and Wilson, 1988). A frequency-following response to the carrier will occur simultaneously with the envelope-following steady-state response when the carrier has a low frequency and the stimulus is loud. Since it is difficult to differentiate from an electrical artefact, it is usually filtered out prior to analysis. Such filtering must be sufficient to prevent aliasing of the carrier response (artefact and frequency-following response) into the frequency region being studied.

Several different auditory responses occur within the 'gamma' band of frequencies (30–80 Hz). As well as the steady-state responses to stimuli repeated at these frequencies, bursts of gamma-band activity may be transiently evoked by stimuli presented more slowly. Galambos (1992) distinguished 'induced' gamma-band activity, which shows trial-to-trial phase variability, from the 'evoked' gamma-band response, which is a brief burst of gamma frequency activity following a transient response and having a stable phase from trial to trial. Induced gamma-band activity has been related to various perceptual and cognitive processes, such as binding (Basar-Eroglu et al, 1996). The evoked gamma-band response occurring at the onset of a sound has a different source from the transient response (Pantev et al, 1993). This would fit with recordings in the rat showing that the two types of response had independent sources in the auditory cortex (Franowicz and Barth, 1995). A recent study distinguished the gamma-band response at the onset of an AM tone from the steady-state response to the ongoing modulation, and showed that the steady-state response developed to its final amplitude within a period of approximately 200 ms that is, perhaps, related to temporal integration (Roß et al, 2002).

Giraud et al (2000) described the activation of different regions of the human brain by AM white noise using functional magnetic resonance imaging. The superior olive, inferior colliculus, medial geniculate, Heschl's gyrus, superior temporal gyrus and parietal lobe were all specifically activated by the AM noise (compared to unmodulated noise). The brainstem regions were more responsive at the higher modulation frequencies, whereas the cortical regions responded best to lower frequencies. Interestingly, Giraud et al did not find much activation at the

modulation frequencies where the electric and magnetic responses are largest (32–64 Hz). The key to this discrepancy probably resides in the synchrony of the neuronal activity. Large electric and magnetic fields occur when the neurons fire synchronously. If the neurons are not synchronous, they may not generate measurable fields at the scalp, even with high levels of activity, as measured by bloodflow.

## Subject variables

### Age
The most striking developmental fact concerning the auditory steady-state responses is that the 40-Hz response cannot be reliably recorded in young infants (Stapells et al, 1988; Maurizi et al, 1990, Levi et al, 1993, 1995). The amplitude of the infant steady-state responses decreases monotonically with increasing rate of stimulation, without showing the normal adult enhancement near 40 Hz (Suzuki and Kobayashi, 1984; Levi et al, 1993). The infant response at 40 Hz is about one-half the response at 10 Hz, whereas the adult response is, on average, 1.5 times larger.

Two factors need to be considered in this respect. First, the immature infant cortex is probably unable to support a sustained rhythmic response at rapid rates. The transient middle-latency response is difficult to record in infants, except at low stimulus rates, and then its latency and morphology are quite different from those of the adult response. Jerger et al (1987) showed that the main positive wave (Pa) response of the transient middle-latency response occurs at about 50 ms in neonates, compared to the 30 ms of adults, and is only clearly recognizable at rates of 2 Hz or less. They pointed out that earlier studies reporting clear responses in infants were probably measuring filter-distorted brainstem responses. If the steady-state response is, indeed, largely the result of superposition of transient responses, then the superposition of these delayed responses could not lead to a prominent response at 40 Hz. Stapells et al (1988) synthesized steady-state responses from transient response waveforms, and demonstrated that steady-state responses may not be recordable even if superposition leads to a recognizable response, since there may be an increased susceptibility to stimulus rate in infants. Second, recordings from an infant are difficult to obtain when the infant is awake, and sleep attenuates the 40-Hz response even in adults. The transient middle-latency response is reduced or absent during delta sleep in infants and young children (McGee et al, 1993; McGee and Kraus, 1996).

It is not clear when the 40-Hz auditory steady-state response develops into adult morphology. The transient middle-latency response may reach the adult pattern by age 12 years (Suzuki and Hirabayashi, 1987; McGee and Kraus, 1996), but there have not been any studies of the steady-state response. The 40-Hz response in infants and children under the age of 6 years is basically a repeated brainstem response (Suzuki and Kobayashi, 1984). Aoyagi et al (1994a) found a general increase in the detectability of the 40-Hz steady-state response from 6 months to 15 years of age. However, the development of the auditory steady-state responses in childhood remains an open question.

Unlike the 40-Hz response, auditory steady-state responses at rates near 80 Hz are reliably recorded in newborns (Rickards et al, 1994; Lins et al, 1996) and in sleeping children (Aoyagi et al, 1993b). On average, the responses in the first few months of life

are between one-third and one-half the size of the adult response, and the thresholds are 10–15 dB higher (Lins et al, 1996). The phases of the responses were quite similar between newborns and adults, except for carrier frequencies of 500 Hz, where the onset phase was about 90° earlier (and the phase delay greater) in the newborn (Lins et al, 1996). Given the latency differences in the auditory brainstem response, one might have expected more significant differences between the phases of the newborn and adult responses. However, latency changes of the order of 1 ms would only show up as a change in phase of about 30° at these stimulus rates, and this might be difficult to measure, given the intersubject variability. Savio et al (2001) found that the amplitude of the response increased over the first year of life, with the responses to higher carrier frequencies showing a greater change.

The auditory steady-state responses do not change significantly with increasing age in adulthood. Johnson et al (1988) found no difference with age in the 40-Hz response, but the young and elderly groups were small. Muchnik et al (1993) compared the 40-Hz response to 500-Hz tone-bursts between 81 young subjects and 20 elderly subjects, and found no age-related changes in the latency or amplitude of the response or the effects of intensity and binaural interaction on these variables. However, if the intensity of the stimulus was measured relative to the thresholds of the subjects (which increased by 9 dB with age, despite the hearing being within normal limits), there was a small age-related increase in the amplitude of the response. This might have been due to some recruitment in the elderly. Boettcher et al (2001) found no difference between young and elderly subjects in the amplitude or phase of the 40-Hz response to amplitude modulation or in the relationship of this response to tonal frequency or modulation depth. In an unpublished experiment, we recorded the steady-state responses to a 1000-Hz tone simultaneously modulated at 3, 43 and 95 Hz and presented at 70 dB SPL to 30 subjects with ages distributed between 20 and 81 years. We found no significant change in the amplitude or phase of any of the responses with age, although the intersubject variability was high, and small changes could have been missed. The slopes of the amplitude regressions were −0.02, +2.6 and −0.5 nV/year for the 3-Hz, 43-Hz and 95-Hz responses, respectively (average amplitudes and standard deviations 360±184, 488±165 and 116±71 nV).

## Sleep and arousal

Sleep has profound effects on the electrical activity of the brain. In general, the effects of sleep are greater for responses that are generated later in the analysis of the auditory information. Given the concept that the steady-state responses at more rapid rates are generated in earlier parts of the auditory system, one might therefore assume that the slower responses are more susceptible to sleep, and this seems to be true to the extent that it has been examined. Sleep significantly reduces the amplitude of the 40-Hz response (Galambos, 1981; Linden et al, 1985; Jerger et al 1986; Rees et al, 1986; Palaskas et al, 1989; Cohen et al, 1991; Aoyagi et al, 1994b; Dobie and Wilson, 1998). On average, the response recorded during sleep is about one-half that recorded during wakefulness. The two studies that looked at sleep stages (Linden et al, 1985; Jerger et al, 1986) did not find any differences, but this may have been due to the small number of subjects for whom data were available from all stages. There is

some suggestion in the data that the amplitude may be smaller in stage 4 sleep than in other stages. This possibility would fit with the results of Pethe et al (2001), who found that the amplitude of the 40-Hz response was inversely proportional to the amount of delta activity in the EEG. Nevertheless, the difference in amplitude among the different sleep stages is less than the difference between sleep and waking.

Several studies have considered the differences in threshold for the 40-Hz response between sleep and waking. Although the steady-state response amplitude decreases during sleep, the background noise (from both the EEG and scalp muscles) may also decrease. In general, thresholds for the 40-Hz response increase by 10–20 dB during sleep (Galambos, 1981; Klein, 1983a; Picton et al, 1987a), although methods extrapolating back to threshold may not show these differences (Linden et al, 1985). Two possibilities must therefore be considered: the threshold for eliciting the response may actually increase, or the threshold may stay the same and the response may be more difficult to recognize during sleep.

Galambos and Makeig (1988) noted changes in the amplitude of the auditory 40-Hz response during drowsiness. As a subject 'dosed off' for one to several minutes, the amplitude of the response decreased to about one-half of what it was in alert wakefulness, and then returned when the subject was 'awake again'. Similar 'minute rhythms' in the 40-Hz response occurred in several different situations, and may indicate a tendency of the brain to fluctuate in its general level of arousal or focus of attention over periods of 30 s to 2 min.

The steady-state responses to stimulus rates more rapid than 70 Hz are much less affected by sleep than the 40-Hz response. Cohen et al (1991) looked at the effects of sleep for stimulus rates between 20 and 180 Hz. They found that the sleep significantly attenuated the response at frequencies below 70 Hz, but not at higher frequencies. These findings were replicated by many others (e.g. Levi et al, 1993; Lins and Picton, 1995).

## Attention

Linden et al (1987) looked at the effects of attention on the auditory 40-Hz response. In a dichotic listening paradigm, a train of tone-bursts was presented to each ear, with the trains differing in rate (37 and 41 Hz) and in tonal frequency (500 and 1000 Hz). The subjects paid attention to one train in order to detect occasional brief changes in frequency, and ignored the other train. Although attention markedly enhanced the transient potentials evoked by the brief changes, the steady-state responses to the ongoing trains were not affected by the direction of attention. These results differ from those reported for the transient middle-latency responses, which can be significantly altered by selective attention (Woldorff et al, 1987).

Makeig and Galambos (1989) used changes in the ongoing auditory steady-state responses to study what was happening in the brain when particular events occurred. The 'complex event-related potential' (CERP) is a time series of complex numbers describing the amplitude and phase at a particular frequency following a discrete event. Changes in the amplitude or frequency of 40-Hz tone-bursts or omissions of a single tone-burst caused changes in the CERP at 40 Hz that persisted beyond the stimulus change. Characteristically, the amplitude decreased for about 200 ms, and then increased. Makeig and Galambos (1989)

suggested that this response might indicate some general process in the brain associated with orientation to novelty. Rohrbaugh et al (1989, 1990) described event-related perturbations in a background 40-Hz response to repeating tones when foreground stimuli occurred. They related the transient decrease in amplitude and latency of the steady-state response to orienting. Rockstroh et al (1996) demonstrated a decrease in the auditory steady-state response when the P300 wave occurred during the detection of an occasional change in the auditory stimulus, and that this reduction was more pronounced when the subject responded by pressing a button rather than by counting. Makeig et al (1996) found that the 40-Hz steady-state auditory response was affected during preparation and execution of simple movements unrelated to the auditory stimuli.

The relationship between the auditory steady-state response and attention thus remains unclear. The 40-Hz response may be affected by attentional changes that are not specific to the information being processed but are more general: changes in alertness or arousal, but not changes in how much information is processed in one selected channel compared to another.

## Audiometric applications

### General comments

An important use of the auditory evoked potentials is to assess hearing—evoked potential audiometry. In many situations, an audiologist needs to measure the hearing of a patient who cannot respond reliably on behavioral testing: infants and young children, emotionally disturbed or cognitively impaired patients, and comatose or anesthetized patients. In these cases, evoked potentials can provide an 'objective' test of hearing, one in which the subject does not have to be subjectively involved. Several aspects of hearing can be tested using the evoked potentials. The simplest test is a screening test that determines whether or not a subject has a hearing loss. The next level measures the thresholds at which sounds of different frequencies can be heard, and constructs an 'audiogram'. The third type of test evaluates how a patient can discriminate different sounds, most importantly those that make up speech. The first two tests—screening and estimating hearing thresholds—depend upon the relationship between the response and the stimulus intensity. This relationship has been extensively studied for auditory steady-state responses at stimulus rates near 40 and 90 Hz. Responses at these rates are the main candidates for audiometric applications.

Three main methods can be used to measure the thresholds for recording the auditory steady-state response. The first method—'threshold bracketing'—involves recording separate responses at several stimulus intensities. The examination starts at an intensity high enough to evoke a response, and then the intensity is decreased until a response can no longer be recognized. The threshold is defined as the lowest intensity where the response can be detected.

A second method involves 'extrapolation'. Responses are recorded at one or more intensities above threshold, and threshold is extrapolated from the amplitude and/or phase of these responses. The extrapolation can be based on normative data for a stimulus measurement, on normative data for the relationship of this measurement to intensity, or on the relationship to intensity shown in the individual data. For example, if an individual's 40-Hz response is recorded at 70 dB HL with an amplitude of 700 nV, and the normative data show that a response with this amplitude is normally recorded at 40 dB above threshold, one might estimate the individual's hearing threshold at 30 dB HL. If a second response was recorded in the same individual at 60 dB HL with an amplitude of 500 nV, one could extrapolate from the recorded 20 nV/dB slope that the response would be absent at an intensity 25 dB lower than 60 dB HL (at 35 dB HL).

A third approach is the 'intensity sweep' technique, which records the response as the intensity is slowly but continuously swept from below to above threshold levels. The result provides the response measurements as a function of intensity. This function can then be introduced into a model that includes a threshold estimate. For example, one can fit the amplitude of the response with a model wherein the amplitude remains stable (at the level of the noise in the recording) at intensities below threshold and then increases linearly with increasing intensity above threshold (Rodriguez et al, 1986). Another approach (illustrated on the right side of Figure 13) is to calculate a regression line between the response amplitudes above the estimated noise level and intensity, and to extrapolate this line back to an estimated threshold at the $x$-intercept of an amplitude–intensity function. Since it records response both above and below threshold, the intensity sweep technique includes aspects of both threshold bracketing and extrapolation.

Most audiometric studies use threshold bracketing. Campbell et al (1977) suggested that threshold could be estimated by linear extrapolation from several suprathreshold measurements, but noted that the intensity functions might be non-linear. Extrapolation from a single recording was used by Kuwada et al (1986), but this was mainly done to show the pattern of hearing loss across different frequencies, and further recordings were then performed to estimate the actual thresholds. Once a threshold at one frequency was set in this way, the thresholds at the other frequencies could be estimated relative to the measured threshold by using the relative size of the response at each of the frequencies. Rodriguez et al (1986) found that the sweep technique was less effective than bracketing thresholds with fixed-intensity recordings.

The main problem associated with the extrapolation and sweep methods is the non-linear relationship between response amplitude and intensity. Lins et al (1995) showed that the amplitude–intensity functions in normal subjects had different slopes for different intensity ranges. This is particularly true for patients with a sensorineural hearing loss. In these patients, the amplitude at intensities more than 20 dB above threshold may be quite normal and may change normally with decreasing intensity. As threshold is approached, the amplitude changes more rapidly, and the response vanishes at a level significantly above that predicted by linearly extrapolating from the measurements at higher intensity. It might be possible to prevent these mistakes by limiting the extrapolation to near-threshold data and eliminating data where the slope changes significantly (see visual studies of Norcia et al (1989)).

Once one has assessed the threshold for the physiologic response, one must then estimate the behavioral threshold. The threshold for recognizing the physiologic response is usually higher than the behavioral threshold, for several reasons. One reason is that the response may be generated by a neuronal

system that is separate from that which determines the perceptual response and that has a higher activation threshold. Another possibility is that the response occurs but is not precisely time-locked to the stimulus when the intensity is near threshold. Since the analysis (averaging and/or frequency transformation) requires time-locking, the response may not be recognizable near threshold. Finally, the response may not be recognizable because of an inadequate signal-to-noise ratio in the recording—if one were able to average forever, perhaps one could record a response at 0 dB SL.

Hearing threshold can be estimated from the threshold for the auditory steady-state response using two main approaches. First, one can just subtract the average difference between hearing thresholds and thresholds for recognizing the steady-state responses as determined in a group of subjects in whom both thresholds are available. For example, one might estimate that hearing threshold is 10 dB below the threshold for recognizing the auditory steady-state response. However, in patients with a sensorineural hearing loss, the steady-state response may be recognized at levels closer to hearing threshold than in subjects with normal hearing. The amplitude of the response grows rapidly as intensity increases above threshold, until the response reaches an amplitude equivalent to that of a normal subject at that sound level (Lenarz et al, 1986). This type of amplitude–intensity function is akin to the loudness–intensity function underlying perceptual recruitment. Amplitude–intensity functions in patients with sensorineural hearing loss have not been examined for the 80–100-Hz responses. However, Dimitrijevic et al (2002) reported that the responses in these patients were larger than those of normal subjects when compared across the same sensation level, which would fit with the hypothesis of a recruitment-like amplitude–intensity function. Given this phenomenon, the relationship between physiologic and behavioral thresholds may have a slope that is greater than 1 in a regression equation with the format $y=mx+b$, where $y$ is the hearing threshold being estimated, $x$ the physiologic threshold, $m$ the slope of the regression line, and $b$ the intercept on the $y$-axis. Rance et al (1995) presented an extensive set of data analyzed in this way. These findings have since been extended over a larger number of subjects without substantial alteration (Rance and Briggs, 2002; Rance and Rickards, 2002). Table 2 compares these and other data recorded using single stimuli to some recorded using the multiple stimulus approach. When reviewing figures in the literature, it is important to be aware of the axes, since some data (e.g. the figures in Rance et al (1995) and in Rance and Briggs (2002)) are plotted with the physiologic rather than behavioral thresholds on the $y$-axis.

Using regression equations to estimate the behavioral thresholds is affected by the level of residual noise in the recording as well as the amount of recruitment. This is illustrated diagrammatically in Figure 17. If the responses are recorded over a briefer period of time and there is more residual noise in the recording (N1 versus N2), the response will not be recognized (and threshold determined) until it reaches a larger amplitude (P1 versus P2). The more slowly growing response in the normal subject will therefore not be recognized until a higher intensity above threshold than in a subject with recruitment. If the levels of noise are very low, the normal subject's response will be recognized at intensities closer to threshold. This interaction probably explains the fact that the slopes of Dimitrijevic et al (2002) and Herdman and Stapells (2003), who recorded over periods of 15 min or more before finally deciding whether a response was present or not, are close to 1, whereas those of Rance et al (1995) and Cone-Wesson et al (2002a), for which the recordings lasted no more than 90 s, are greater than 1.

At high intensities it is easy to recognize a response in most subjects (even those with a hearing loss) quite quickly (e.g. a minute or two). Because of recruitment, it is relatively easy to demonstrate near-threshold responses in a short time in patients with sensorineural hearing loss. It is only when you get to near-threshold levels in subjects with mild or no hearing loss that longer times (e.g. 15 min) are required to demonstrate a response or to prove that it is not there. If one uses brief recordings in these subjects, the physiologic thresholds will be higher than if one uses longer times, and the physiologic–behavioral difference will be higher. This effect may be compensated for by using a regression equation that relates the behavioral thresholds to physiologic thresholds. For example, in the regression equation of Rance et al (1995) for the 500-Hz carrier, a physiologic threshold at 40 dB HL would give an estimated threshold of −1 dB HL, whereas a physiologic threshold of 90 dB HL would

**Table 2.** Relationship between hearing threshold and physiologic thresholds

| Frequency (Hz) | Rance et al (1995) | | Cone-Wesson et al (2002a) | | Rance and Rickards (2002) | | Dimitrijevic et al (2002) | | Herdman and Stapells (2003) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | m | b | m | b | m | b | m | b | m | b |
| 500 | 1.30 | −40 | 1.39 | −35 | 1.39 | −49 | 0.88 | −9 | 0.77 | −6 |
| 1000 | 1.18 | −26 | 1.24 | −15 | 1.35 | −44 | 0.92 | −1 | 0.91 | −4 |
| 2000 | 1.05 | −19 | 0.94 | +5 | 1.28 | −36 | 0.89 | 0 | 0.92 | −6 |
| 4000 | 1.19 | −24 | 1.34 | −24 | 1.31 | −39 | 0.99 | −8 | 1.04 | −5 |

This table describes the regression results relating to the physiologic thresholds of the auditory steady-state responses at 70–100 Hz to hearing threshold. The relationship has the format $y=mx+b$ where $y$ is the hearing threshold (in dB HL), $x$ is the threshold for recognizing an auditory steady-state response, and $b$ is the $y$-intercept. The Rance et al (1995) data are based on evaluations of 60 subjects, although the number of points in each regression is less than 60, because of missing data. The data in the Cone-Wesson et al (2002a) and Rance and Rickards (2002) studies include some of the data from the Rance et al (1995) study. The numbers of points for the regressions vary from 18 to 78 for the Cone-Wesson study and from 137 to 260 for the Rance and Rickards study. The Dimitrijevic et al (2002) data are based on 87 points at each frequency (from 46 subjects since, in most of the subjects, both ears were tested). The Herdman and Stapells data are based on between 26 and 29 points for each frequency. The Rance et al (1995), Rance and Rickards (2002) and Cone-Wesson et al (2002a) studies presented stimuli singly, whereas the Dimitrijevic et al (2002) and Herdman and Stapells (2003) studies presented multiple stimuli simultaneously. As discussed in the text, the other major difference between the studies was the time taken to recognize responses.
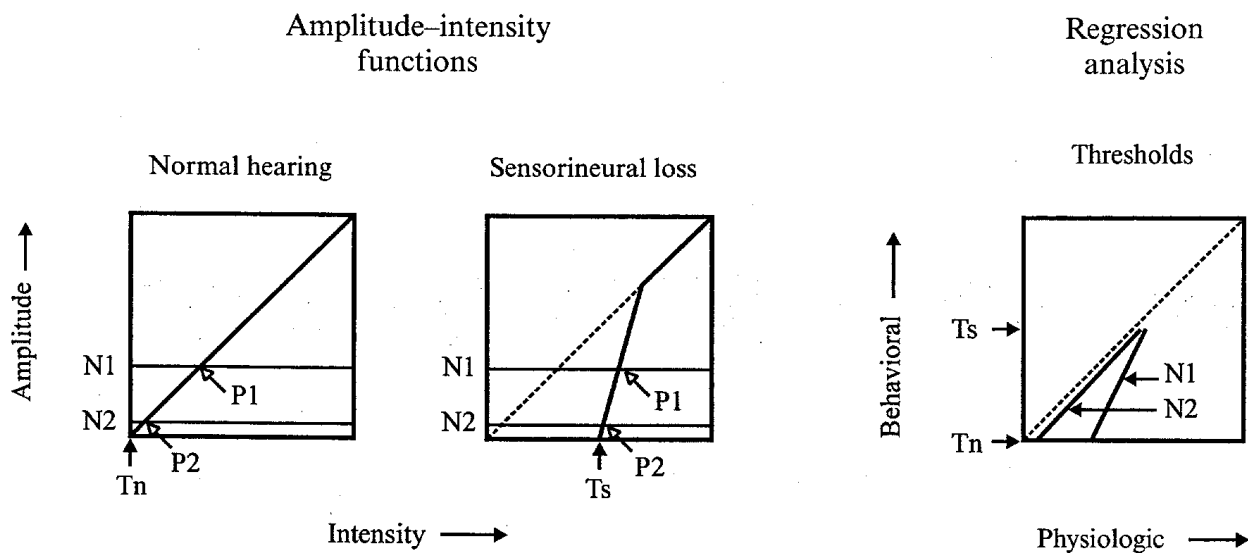
**Figure 17.** Prediction of behavioral thresholds using regression analysis. This figure illustrates diagrammatically some issues concerning the use of regression estimates to predict behavioral thresholds from physiologic thresholds. On the left side of the figure are shown two representative amplitude–intensity functions, one in a subject with normal hearing, and one in a subject with sensorineural hearing loss. The behavioral thresholds are indicated by Tn and Ts for the two subjects, and for simplicity the threshold for recognizing the physiologic response is made equal to the behavioral threshold when the noise in the recording has been attenuated to zero. Physiologic thresholds (P1 and P2) are measured after a short recording time when the residual noise in the recording is N1, and after a longer time when the level of noise is N2. The relationship between the physiologic and behavioral thresholds is analyzed using regression, and the regression lines are plotted in the graph on the right side of the figure. The slope of the regression line is larger when the physiologic thresholds are measured when the noise levels are higher after a shorter recording time (N1 versus N2).

give an estimated threshold of 77 dB HL. However, using regression makes one susceptible to two sources of variance: that obtained from measuring the physiologic threshold, and that inherent in the regression process. Particular problems will occur in patients with conductive hearing loss, or other hearing-impaired patients who do not show large responses at near-threshold levels. In these patients, the physiologic–behavioral differences will be similar to those in normal subjects, but the thresholds will be high, and these will not be sufficiently corrected.

It is essential to clarify the term 'threshold'. We suggest 'physiologic thresholds' to describe the lowest intensity at which an auditory steady-state response is recognized, and 'predicted behavioral thresholds' if one is predicting behavioral thresholds from these physiologic thresholds. Predicted behavioral thresholds would have to be further qualified, since the prediction could derive from a regression equation or from a simple subtraction of the mean difference between physiologic and behavioral thresholds.

*Audiometry using 40-Hz responses*
Galambos et al (1981) were the first to estimate hearing thresholds using the steady-state responses to stimuli presented at rates near 40 Hz. Campbell et al (1977) had presented stimuli at rates as high as 32 Hz, but they had estimated thresholds using the lower rates. Galambos et al (1981) had found that five of seven subjects showed a recognizable response to a 500-Hz tone-burst at 5 dB SL, and the other two at 10 dB SL. Many subsequent studies confirmed that the 40-Hz response can be recorded at intensities close to threshold in both normal and hearing-impaired subjects. Table 3 summarizes some data from

studies that provided threshold estimates at multiple frequencies, and does not include early studies which estimated thresholds only at 250 and/or 500 Hz (Kileny and Shea, 1986; Barajas et al, 1988).

In general, Table 3 shows that physiologic thresholds for the 40-Hz response are about 10 dB above hearing thresholds. However, given a standard deviation of about 10 dB, about one in 20 subjects will not show recognizable responses at 30 dB above hearing thresholds. The one outlier study (Milford and Birchall, 1989) probably recorded for too short a time (about 25 s at each intensity) to recognize small responses.

Two main problems occur when using the 40-Hz steady-state response to assess thresholds. First, the response is difficult to recognize in young infants (see 'Age', above), who are those subjects in whom objective audiometry is most important. Second, the amplitude of the 40-Hz response in a sleeping subject is generally about one-half the amplitude of the response in a waking subject (see 'Sleep and arousal', above). Several initial studies of the 40-Hz response remarked that thresholds were elevated by 15–40 dB when the recordings were performed in sleep (Szyfter et al, 1984; Kankkunen and Rosenhall, 1985). This might explain the results of Barajas et al (1988), who reported thresholds for 500-Hz tone-bursts that averaged 25 dB above hearing thresholds, since many of their subjects were asleep.

*Audiometry using 80-Hz responses*
The auditory steady-state responses evoked by stimuli with modulation rates higher than 70 Hz have several advantages over the 40-Hz response. First, they are much less affected by sleep (see 'Sleep and arousal', above). Second, these rapid

**Table 3.** Threshold estimation using 40-Hz auditory steady-state responses

| Study | Subjects[a] | Stimuli[b] | Analysis[c] | Physiologic–behavioral thresholds (dB) | | | |
|---|---|---|---|---|---|---|---|
| | | | | 500 | 1000 | 2000 | 4000 |
| Klein (1983a) | 30 N | TB | V | 16±10 | 14±7 | 16±7 | 19±7 |
| Szyfter et al (1984) | 31 N | TB | V | 15±9 | 13±7 | | |
| Dauman et al (1984) | 30 H | TB | V | 11±10 | 9±10 | | |
| Lynn et al (1984) | 40 H | TB | V | −2±12 | 8±11 | | |
| Sammeth and Barry (1985) | 16 N | TB | V | 9±7 | 10±10 | 9±5 | 16±7 |
| Kankkunen and Rosenhall (1985) | 20 M | TB | V | 8±11 | 5±9 | 4±7 | 3±8 |
| Rodriguéz et al (1986) | 15 N | TB | T | 3±8 | | 20±9 | |
| | 10 H | TB | T | 4±10 | | 5±10 | |
| Stapells et al (1987) | 6 M | TB | P | 1±3 | | 2±4 | |
| Milford and Birchall (1989) | 22 N | TB | V | | 27±10 | 23±15 | 16±15 |
| Chambers and Meyer (1993) | 10 M | AM | S | 1±9 | 2±8 | | |
| Aoyagi et al (1993b) | 15 N | AM | P | 11±10 | 11±11 | 13±10 | 18±12 |
| | 18 H | AM | P | 8±7 | 9±6 | 13±8 | 12±6 |

[a] Subjects: N, normal; H, hearing-impaired; M, a mixture of normal and hearing-impaired subjects.
[b] Stimuli: TB, tone-burst; AM, is amplitude-modulated tone.
[c] Analysis: V, visual evaluation of time waveform; S, visual evaluation of spectrum; P, phase coherence; T, Hotelling's $T^2$ test.
For many of the studies, the standard deviation has been estimated from the data plotted in the figures, as 1.25 times the mean absolute difference, or as one-third the range.

responses are easily recorded in infants (see 'Age', above). A third advantage is that responses can be recorded to multiple stimuli presented simultaneously without significant loss in amplitude (see 'Multiple simultaneous stimuli', above).

Over the past several years, many studies have demonstrated that the steady-state responses at modulation frequencies of 75–100 Hz can provide reliable estimates of hearing threshold. Table 4 summarizes the results in adults and older children. In general, the 80-Hz responses can be recognized down to 15 dB above hearing thresholds, and the variability is such that one in 20 subjects will show thresholds that are more than 35 dB higher than hearing threshold. Perez-Abalo et al (2001) calculate that only 3% of the thresholds estimated in hearing-impaired children will differ from the actual hearing threshold by more than 40 dB. In Table 4, the data obtained by Aoyagi et al (1994c) in normal adults are aberrant. It is difficult to understand why these thresholds are so elevated, since the subjects were sleeping, the averaging was carried out for 3 min at each intensity, and the phase coherence detection procedure was both objective and efficient. However, we have found that at near-threshold levels, we often need recording times of 10 min or more to recognize the small responses. Herdman and Stapells (2001) recorded responses for up to 13 min (48 sweeps of 16 s) when evaluating their near-threshold responses. Dimitrijevic et al (2002) recorded for up to 17 min.

Cone-Wesson et al (2002a) compared thresholds in sleeping subjects for auditory brainstem responses to tone-bursts, and for the auditory steady-state responses to mixed-modulation tones with the modulation at different rates. The average thresholds for the auditory brainstem responses were 13 and 7 dB SL at 500 and 4000 Hz, respectively. The thresholds for the auditory steady-state responses for modulation rates of 41 Hz were 23 and 33 dB SL. At higher modulation rates (74 Hz for 500 Hz, and 95 Hz for 4000 Hz), the thresholds were 40 and 16 dB SL.

The thresholds for the auditory steady-state responses were higher than the auditory brainstem responses and higher than other reported thresholds for the steady-state responses (Tables 3 and 4), probably because of the brief recording period (maximum 100 s). However, the comparison across the different modulation rates is instructive, since it was performed in sleeping subjects and using recordings of similar duration. The thresholds for the 500-Hz carrier were better at 41 Hz than at 74 Hz, whereas the thresholds for the 4000-Hz carrier were better at the higher modulation rate. In young children with normal hearing and in those with hearing impairment, thresholds estimated using auditory brainstem responses are highly correlated with those estimated using auditory steady-state responses (Cone-Wesson et al, 2002a; Vander Werff et al, 2002).

Table 5 summarizes the results obtained with the 80-Hz responses in infants. The thresholds in infants are higher than those obtained in children or adults. However, two factors need to be considered when evaluating these data. Some of the tabulated data were not obtained in a sound-attenuated room, and may be elevated by 10 or 15 dB over what might be expected in a less noisy environment. Furthermore, some of the thresholds were estimated on the basis of recordings lasting for only a few minutes (for example, Rickards et al (1994) used less than 90 s, and Levi et al (1993) used recordings that lasted between 100 and 200 s). Taking these effects into account, the thresholds in infants are elevated by 10–20 dB compared to thresholds in adults (or children older than 1 year), and there is a decrease in thresholds over the first year of life. Thresholds at the lower frequencies (500 Hz) are more elevated than those at at higher frequencies. This is compatible with the data of Cone-Wesson et al (2002b), who found that the pass rate for a rapid screening test of newborn infants was significantly higher for a carrier frequency of 2000 Hz at 49 dB SPL (86%) than for a

carrier frequency of 500 Hz at 62 dB SPL (65%). Savio et al (2001) reported that thresholds decrease over the first year of life, with the decrease being larger for the higher frequencies (e.g. 7 dB at 500 Hz, and 14 dB at 4000 Hz). This is similar to the maturation of the auditory brainstem response. These phenomena may be related to maturational processes in both the cochlea and the nervous system. More extensive information on the normal infant response is clearly needed.

Rance et al (1998) used auditory steady-state responses to assess ears where there is very little residual hearing. Auditory brainstem responses are often absent in these patients. Clicks (or brief tones) which have a high peak SPL at their normal hearing threshold cannot be presented at sufficiently high intensities to evoke a response, due to concerns about peak SPL levels. However, relationships between the intensity of the stimulus and hearing threshold are not simple. The intensity of a click 95 dB

**Table 4.** Threshold estimation using 80-Hz auditory steady-state responses

| Study | Subjects[a] | Stimuli[b] | Analysis[c] | Physiologic–behavioral thresholds (dB) | | | |
|---|---|---|---|---|---|---|---|
| | | | | 500 | 1000 | 2000 | 4000 |
| Aoyagi et al (1994c) | 20 N | AM | P | 34±15 | 29±14 | 30±15 | 9±14 |
| Rance et al (1995) | 60 M | MM | P | 20±7 | 13±6 | 16±5 | 10±4 |
| Lins et al (1996) | 15 N | AM | F | 14±11 | 12±11 | 11±8 | 13±11 |
| Picton et al (1998) | 10 N | AMF | F | 21±9 | 26±13 | 18±13 | 20±10 |
| | 35 A | AMF | F | 17±8 | 13±7 | 13±7 | 16±9 |
| Rance et al (1998) | 108 H | MM | P | 6±7 | 4±6 | 3±6 | 6±7 |
| Aoyagi et al (1999) | 125 HC | AM | P | 15±15 | 4±13 | 9±16 | 2±14 |
| Herdman and Stapells (2001) | 10 N | AM | F | 14±10 | 8±7 | 8±9 | 15±9 |
| Perez-Abalo et al (2001) | 40 N | AM | F | 12±11 | 13±9 | 10±10 | 13±10 |
| | 43 HC | AM | F | 13±15 | 7±15 | 5±15 | 5±16 |
| Dimitrijevic et al (2002) | 45 M | MM | F | 14±11 | 5±9 | 5±9 | 9±10 |
| Herdman and Stapells (2003) | 29 H | AM | F | 14±13 | 8±9 | 10±10 | 3±10 |
| Rance and Briggs (2002) | 184 HC | MM | P | 6±9 | 6±7 | 4±8 | 3±11 |

[a] Subjects: N, normal; H, hearing-impaired; M, mixture of normal and hearing-impaired subjects; C, children; HC, hearing-impaired children; A, using hearing aids.
[b] Stimuli: AM, amplitude-modulated tone; AMF, amplitude-modulated tones presented in free field; MM, mixed-modulation tone.
[c] Analysis: P, phase coherence; F, F-test of spectral data. Studies with phase coherence used single stimuli, and studies with the F-test used multiple stimuli.

For the Rance et al (1998) data, the difference has been estimated from the regression (Table 1) at 70 dB HL and the SD from the overall standard errors of the regressions. For the Aoyagi et al (1999) data, the means and standard deviations have been estimated from the regression scatter-plots (except for 1000 Hz, where the data were presented).

**Table 5.** Thresholds for 80-Hz auditory steady-state responses in normal infants

| Study | Subjects[a] | Stimuli[b] | Analysis[c] | Physiologic–behavioral thresholds (dB–HL)[d] | | | |
|---|---|---|---|---|---|---|---|
| | | | | 500 | 1000 | 2000 | 4000 |
| Rickards et al (1994) | 337 | MM | P | 41±10 | 24±9 | | 35±11 |
| Levi et al (1995) | 35 | AM | M | 45±16 | 49±11 | 60±15 | |
| Lins et al (1996) | 21 | AM | F | 34±13 | 20±10 | 18±8 | 24±10 |
| | 30 | AM | F | 47±12 | 34±14 | 31±12 | 33±12 |
| Savio et al (2001) | 64 | AM | F | 56±12 (16) | 52±12 (22) | 50±12 (19) | 50±13 (23) |
| Cone-Wesson et al (2002c) | 85 | MM | P | 34±8 | 25±10 | 21±10 | 31±12 |

[a] Subjects: the infants were less than 1 year old. Most infants were examined within the first 6 months. All babies in the Rickards study were evaluated within the first week after birth. The babies in the Levi study were 1 month old. For the Savio et al data, the tabulated thresholds are for babies in the first month of age. For the Rickards and Cone-Wesson studies, not all carrier frequencies were evaluated in all infants.
[b] Stimuli: AM, amplitude-modulated tone; MM, mixed-modulation tone.
[c] Analysis: P, phase coherence; M, magnitude-squared coherence; F, F-test of spectral data. Data analyzed with coherence measures were for single stimuli, and data analyzed with the F-test were for multiple stimuli. The single-stimulus recordings usually lasted for less than 2 min whereas the multiple stimulus recordings lasted for 8–13 min for near-threshold intensities.
[d] For these data, there are no behavioral thresholds. Therefore, intensities reported in SPL have been converted to HL using the normative values of Herdman and Stapells (2001): 11, 9, 8 and 7 dB SPL for 500, 1000, 2000 and 4000 Hz respectively. Levi et al (1995) reported thresholds of 18, 7 and 11 dB SPL at 500, 1000 and 2000 Hz. For the Rickards et al (1994) data, the 1500-Hz thresholds have been tabulated between 1000 and 2000 Hz. In the Lins et al (1996) paper, the two groups of babies were tested in two different places (21 in Ottawa and 30 in Cuba). The data of Savio et al (2001) are given first in HL as described above, and then (in parentheses) in dB nHL as reported in their paper (SPL level minus the average adult behavioral thresholds in their recording environment). Of these tabulated data, only the Levi et al (1995) data, the Lins et al (1996) Ottawa data (21 B) and the Cone-Wesson et al (2002c) data were recorded in a sound proof room with acceptable acoustic noise levels.

above the normal threshold for that click has a peak intensity of 135 dB pSPL, which is the same as for a 132 dB rms SPL continuous tone. It is not clear how behavioral and physiologic thresholds are related at such high intensities. Furthermore, caution must be exercised in presenting high-intensity continuous tones for prolonged periods, since noise-induced damage to the hair cells might result if the neural response is not rapidly recognized.

*Frequency specificity*

The frequency specificity of a stimulus–response pairing assesses how much better the response is evoked by a stimulus of a particular tonal frequency than by a stimulus of another frequency. Three levels need to be considered in evaluating the frequency specificity of a stimulus–response technique: acoustic specificity, place specificity, and neuronal specificity (Stapells et al, 1994; Herdman et al, 2002b).

The acoustic specificity of a stimulus measures how well the energy within the stimulus is concentrated within a particular set of frequencies in the spectrum. Continuous pure tones have maximum acoustic specificity. Problems of acoustic specificity arise when brief tone-bursts (or tone-bursts with brief onsets) are used to elicit responses, since these stimuli have energy in the spectrum at frequencies away from their nominal frequency— 'spectral splatter'. A classic problem involving reduced acoustic specificity is when a brief high-frequency tone-burst evokes a response (behavioral or physiologic) from an individual with a severe high-frequency hearing loss but near-normal hearing at lower frequencies. The response is mediated by the low-frequency energy in the spectrum of the brief tone-burst (Picton, 1978). Recording transient evoked potentials requires a compromise between stimuli that are sufficiently frequency-specific to estimate pure-tone thresholds, and stimuli that have sufficiently rapid onsets to evoke easily recognizable responses. More rapid onsets lead to larger responses but also decrease the frequency specificity of the stimuli. The spectral purity of a brief tone-burst can be maximized by shaping the envelope with specialized functions (Gorga and Thornton, 1989). Although theoretically better, these tone-bursts do not have greater frequency specificity when evaluated experimentally (Oates and Stapells, 1997). Sinusoidally amplitude-modulated stimuli are acoustically very frequency-specific, with energy only at the carrier frequency and at two sidebands separated from the carrier frequency by the modulation frequency (Figure 8).

Place specificity measures the extent of the basilar membrane that is activated by the stimulus. The traveling wave causes an activation pattern with a more apical maximum for tones of lower frequency, and a greater basal than apical spread of the activation from this maximal point. The point of maximal activation can be considered the specific place for that tonal frequency. The spread of activation to regions of the basilar membrane other than this specific place can make it difficult to determine which place is initiating the response. A classic problem involving reduced place specificity occurs when a low-frequency tone is heard by an individual with no functioning hair cells beyond the first turn of the cochlea. The response is initiated through regions of the cochlea that respond best (are place-specific) to higher-frequency tones. Place specificity is measured by analyzing the extent of the cochlea that is activated by the stimulus. One method uses masking with high-pass noise to isolate responding regions of the cochlea. Presenting the stimuli in high-pass noise with decreasing cut-off frequencies, and then sequentially subtracting the responses, can provide 'derived responses' that measure the activation of the cochlear regions between the two cut-off frequencies of the high-pass noise.

The place-specificity of the 40-Hz steady-state response evoked by tone-bursts was studied by Klein (1983b). He used tone-on-tone masking to map out the tuning curve of the response by measuring the intensity of the masking tone needed to reduce the amplitude of the steady-state response by half. The tuning curves showed good frequency specificity (less than an octave wide 20 dB above the tip of the curve) at frequencies of 500–4000 Hz, and the frequency specificity was preserved up to 80 dB SPL. High-pass noise masking (Griffiths and Chambers, 1991) showed that the 50-Hz steady-state response evoked by amplitude modulation of a 70 dB SPL tone is mediated by a narrow range of activation on the cochlea—an octave wide or less.

The place-specificity of the 80-Hz auditory steady-state response at 60 dB SPL has been evaluated by Herdman et al (2002b) using high-pass masking. They found no significant differences between simultaneously presenting four stimuli separated by octaves and presenting single stimuli. The half-amplitude bandwidth for the activation pattern of the derived response was, on average, between 1 and 1.25 octaves for carrier frequencies between 500 and 4000 Hz. This bandwidth is slightly narrower than that for the auditory brainstem responses to brief tone-bursts, but not as much as might be expected from the increased acoustic specificity of AM tones compared to brief tone-bursts. Place-specificity in the cochlea is probably the limiting factor. One cannot be more frequency-specific than the cochlear filter.

Neuronal specificity should be considered when discussing the frequency specificity of auditory evoked potentials. In a normal cochlea responding to single tones, the tuning curves of the auditory nerve fibers and the hair cells are similar. The frequency-specificity characteristics of the primary auditory neurons are preserved in a proportion of central neurons, although other central neurons have much broader tuning curves. When there is a cochlear hearing loss, the tuning curves of auditory nerve fibers are distorted, exhibiting diminished tips and increased low-frequency tails. These fibers may respond better to frequencies much lower than their characteristic frequency (or from what might be expected from their place on the basilar membrane). This decrease in the neuronal specificity may lead to responses that are not very frequency-specific.

When multiple stimuli occur together, processes such as suppression (in the cochlea) and lateral inhibition (in the central nervous system) may alter the neuronal specificity of the responses. The results of Herdman et al (2002b) suggest that these processes do not significantly affect the responses when the carrier frequencies are separated by one octave and when the stimuli are presented at moderate intensity levels. At higher intensities, significant interactions between all stimuli may occur (John et al, 1998). Even at moderate intensities, there may be some attenuation of the lower-frequency responses by concomitant high-frequency sounds (John et al, 1998; Dolphin and Mountain, 1993). In a situation of high-frequency hearing loss, lower-frequency (i.e. 500-Hz) AM tones may also mask the auditory steady-state responses to high-frequency (i.e. 4000-Hz) stimuli (Picton et al, 1998).

## Masking

Broadband noise provides another way to evaluate the specificity of a neuronal response. Contralateral noise has no significant effect on the auditory steady-state responses when the modulation frequencies are 80–110 Hz (Lins et al, 1995). However, contralateral noise at levels much below those that would result in cross-hearing significantly decreases the amplitude of the 40-Hz response (Galambos and Makeig, 1992a,b). These results indicate that neurons generating the 40-Hz response can be activated by sounds in either ear, whereas those generating the 80-Hz response are mainly monaural. This would fit with studies of binaural interaction which show that at 80 Hz the binaural response is essentially the sum of the two monaural responses (Lins et al, 1995), whereas at 40 Hz the binaural response is much smaller than the sum of the monaural responses (Picton et al, 1985). Ipsilateral noise has complex effects on the 40-Hz responses (Galambos and Makeig, 1992a,b). At low levels, there is some increase in the response amplitude, whereas at higher levels the response has a decreased amplitude and shorter latency. The unexpected latency change might possibly be caused by the noise masking a cortical generator and making the brainstem generator more prominent in the response.

## Bone conduction

Two techniques are available for evaluating bone conduction thresholds with the auditory steady-state responses. The first technique presents the stimuli directly through a bone conduction transducer positioned either on one of the mastoids or on the forehead. Thresholds obtained using this technique for the 80–100-Hz responses are approximately 10 dB above the behavioral thresholds for the same bone-conducted tones, with approximately the same intersubject variability as for air-conducted stimuli (Lins et al, 1996; Dimitrijevic et al, 2002). A disadvantage of this technique is that electrical artefacts at the modulation frequency may be generated by the passage of the stimuli through the vibrator. These artefacts might be ruled out if the responses are eliminated when the stimuli are presented in masking noise, since this should not affect artefacts (Dimitrijevic et al, 2002). However, ruling out the artefacts requires extra recordings, and if the artefacts are not ruled out, the bone conduction thresholds are uncertain.

Cone-Wesson et al (2002c) used the 'sensorineural acuity level' technique to estimate bone conduction thresholds. This procedure evaluates the amount of masking noise presented by bone conduction that is needed to mask an air-conducted stimulus that is just above threshold for that subject. For a purely conductive hearing loss, the amount is equivalent to the levels that occur for normal-hearing subjects. Noise above the normal level needed to mask in the cochlea elevates the air-conducted threshold, which has been already elevated by the conductive loss. For a purely sensorineural hearing loss, the amount of bone conduction noise necessary to elevate the air-conducted threshold is higher than in normal subjects, since the threshold for any masking in the cochlea is elevated by the sensorineural loss. The amount by which the noise has to be increased represents an evaluation of the bone conduction threshold or the sensorineural acuity level. Cone-Wesson et al (2002c) found that this technique successfully categorized infants with and without a sensorineural hearing loss as determined by tympanometry and auditory brainstem responses.

This approach to estimating the bone conduction thresholds has the advantage that any artefacts caused by the bone vibrator are irrelevant to recording the response (to air-conducted stimuli), although they may increase the noise level of the recording. A disadvantage is that the threshold is based on two threshold estimations (with and without masking), and thus is more variable than a directly estimated threshold.

## Objective evaluation of suprathreshold hearing

Although determining hearing thresholds at particular tonal frequencies (obtaining an audiogram) is an essential part of the audiometric assessment, measuring the ability of a subject to discriminate suprathreshold sounds also provides meaningful information. An inability to understand speech is the usual presenting symptom of a patient with a hearing problem, and an improvement in the ability to discriminate speech sounds is the way in which treatment (with surgery, hearing aids, or cochlear implants) is demonstrated and monitored. An objective technique to measure suprathreshold hearing would therefore be very helpful in the investigation of a patient's ability to understand speech, in the selection and monitoring of hearing aids, and in studying disorders of auditory perception.

Stimuli with independent modulation of amplitude and frequency (IAFM; see 'frequency modulation and mixed modulation', above) may become useful in assessing how well the brain can detect changes in amplitude and frequency. We have hypothesized that the responses to several such stimuli may allow us to assess the ability of the cochlea and brainstem to discriminate sounds that are important for speech discrimination. Dimitrijevic et al (2001) showed that the number of significant responses to four carrier frequencies (eight modulations) spanning 500–4000 Hz varied with the word discrimination score when the stimuli (IAFM or words) were presented at different intensities. However, although the measurements were significantly related, it was not possible to predict the word recognition score accurately in individual subjects.

The general idea of this approach is that the IAFM responses provide an index of the auditory information that has been properly processed through the brainstem for presentation to the cortex. These measures should therefore correlate with word discrimination changes caused by acoustic manipulations (intensity, masking, amplification), conductive or cochlear hearing losses, auditory nerve dysfunction (neuroma, neuropathy), and brainstem disorders. They should not correlate when word discrimination deficits are caused by cortical dysfunction. Steady-state responses recorded using lower modulation rates may correlate more closely with cortical processing.

The steady-state responses may also demonstrate the ability of the auditory system to process rapid temporal changes. The ability of the auditory system to follow rapid acoustic changes is assessed psychophysically by means of temporal modulation transfer functions (Viemeister, 1979), or the threshold for detecting gaps in a sound (Moore and Glasberg, 1988). The effect of modulation frequency on the auditory steady-state responses may be related to the temporal modulation transfer function. One might be able to compare the response amplitude to normative values, or assess the detectability, rather than simply the amplitude, of the response. However, the relationship is not simple, since the psychophysical function does not have a peak near 40 Hz (Figure 10).

## Hearing aids and cochlear implants

At present, the most important role for the auditory steady-state responses in fitting hearing aids or in determining whether a child should have a cochlear implant is in providing an accurate and objective assessment of hearing at different frequencies (Rance et al, 1993; Picton et al, 2002a,b). The fitting of a hearing aid is largely based on an accurate audiogram, and this is difficult to obtain in young infants without using physiologic measurements. Deciding on a cochlear implant likewise requires some accurate assessment of the severity of the hearing loss.

Auditory steady-state responses may also be helpful in evaluating the actual functioning of hearing aids, since they can be evoked by AM tones, which are frequency-specific and stable over time. Compared to transient stimuli, these stimuli are much less likely to be distorted by amplification in either a sound-field speaker or a hearing aid. In general, hearing aids handle rapidly changing acoustic stimuli differently from more continuous stimuli such as speech, and it is difficult to predict the steady-state characteristics of hearing aids from onset responses (Gorga et al, 1987).

Picton et al (1998) showed that auditory steady-state responses to AM tones with modulation frequencies between 80 and 105 Hz can be recorded when multiple stimuli are presented simultaneously through a sound-field speaker and amplified using a hearing aid. Responses were recorded down to intensities close to the behavioral thresholds for sounds in the aided condition. It is therefore possible to assess aided thresholds. However, the amplification of a hearing aid is usually not linear across stimulus intensities, with higher amplification occurring at lower intensities. This may lead to discrepancies in the physiologic measurement of gain.

The role that these studies might play in fitting hearing aids is not clear. The gain of a hearing aid can be measured using acoustic measurements in the ear canal. If the threshold of the subject is known from behavioral or physiologic methods, the gain of the hearing aid can be adjusted to compensate for the measured hearing loss without the need for assessing aided thresholds. The most important use for the steady-state responses may therefore be to obtain accurate unaided thresholds. The actual gain of the aid could then be measured and adjusted on the basis of within-the-canal acoustic measurements. Nevertheless, demonstrating that the hearing aid is causing sounds to activate responses in the brain at intensities where there was no response without the aid is an important confirmation of the benefit of the aid. This is essential in patients who do not have clear or reliable thresholds (either behavioral or physiologic) without aids.

A more important application of the steady-state responses in fitting hearing aids would be to assess suprathreshold hearing. Even though the gain of an aid can be adjusted on the basis of within-the-canal acoustic measurements, one still does not know how the aided sound is processed within the brain. Two major interrelated questions are whether the amplified sounds are optimally discriminable, and whether their loudness is comfortable. A measure of uncomfortable loudness level is an essential piece of information in fitting a hearing aid, since it determines the maximum amplification that should be provided by the aid. An accurate measure of this level is not available in young children. Parker and O'Dwyer (1998) suggested that the 40-Hz response to a 400-Hz tone might be helpful in this regard.

They found that they could predict the uncomfortable loudness level from the change of phase with intensity in responses recorded at intensities below the uncomfortable loudness levels.

Some measure of the discriminability of sounds presented through a hearing aid is essential for demonstrating the benefit of a hearing aid. The sounds might be adequately amplified so that they are above the threshold for detection, and yet might not be discriminable due to distortion in the aid or in the physiologic processing. Adequate word recognition scores provide a simple way of demonstrating discrimination in subjects who can process words, but this is not possible in young subjects. We have studied the responses to multiple AM and FM stimuli presented simultaneously to patients using hearing aids (Picton et al, 2002b). The number of recognizable responses varied with the word recognition scores. However, the relationship is not exact, and the technique needs to be significantly improved before we can make a reasonable prediction of word discrimination from these responses.

## Time and accuracy in evoked potential audiometry

The use of the auditory steady-state responses to assess hearing is not a completely exact science. Responses need to be distinguished from noise, and the accuracy of this procedure will vary with the size of the response, the amount of electrical noise in the recording, and the time taken to reduce this noise by averaging or increasing the sweep of the FFT analysis. There has been a general tendency to try to estimate thresholds as rapidly as possible. However, the standard deviations of the physiologic-behavioral differences (e.g. Table 4) are generally near 10 dB. This means that the behavioral threshold in one in 20 estimations will be at least 20 dB different from the physiologic thresholds, even after the normal physiologic–behavioral difference has been removed by subtraction or regression. In one in 100 estimations, the differences will be more than ±30 dB. Allowing more time for the recordings should lead to greater accuracy.

The multiple stimulus technique is therefore quite appealing. Since responses to multiple stimuli can be recorded in the same time that it takes to record one, one can record for a longer time and obtain higher signal-to-noise ratios for detecting the responses, and thus more accurate estimations of behavioral thresholds. However, the time-saving is not as great as one might initially assume. Recording responses to eight stimuli simultaneously does not make threshold estimation eight times faster (John et al, 2002b). In normal subjects, the difference in amplitude across carrier frequency causes some responses to become significant before others. This increases the time needed when assessing thresholds, since the recording period at each intensity has to be extended so that the smallest of the multiple responses becomes significantly larger than noise. Another issue that arises when performing threshold evaluations is that it takes longer to determine that a response is absent than it does to recognize that a response is present. The decision that a response is not present usually requires that the response is not recognizably different from the background noise after this noise has been reduced to a criterion level. If the subject has a sloping audiogram, there are several intensities at which an investigator must decide that at least one response is absent. Recordings at these intensities need to be continued until the criterion noise level is reached (rather than until the other responses are

recognized). These factors attenuate the advantage of the multiple stimulus approach. However, we have found that using four stimuli in each ear (eight total) increases the speed of the assessment by two to three times. Other approaches that dynamically adjust the intensities of the different stimuli on the basis of whether a response has been recognized may increase this advantage (John et al, 2002b).

Any threshold estimation must be qualified by the noise level at which a response is judged to be absent. These levels will determine both the accuracy of the thresholds and the overall time taken to estimate them. Both Dimitrijevic et al (2002) and Herdman and Stapells (2002) used the criterion that a response at the 80–100-Hz rate was absent when there was no significant response with a mean noise level less than 10 nV. Using the F-test, this would mean that a response with amplitudes less than 17 nV might be missed (John et al, 2002b). Given such a criterion, John et al estimated that the time taken to measure a four-frequency audiogram for a subject with a sloping audiogram and different thresholds at each frequency was about 83.5 min. For other audiometric patterns where thresholds were the same at two or more frequencies, the time would be less. Herdman and Stapells (2003) found the actual time to measure a four-frequency audiogram in a group of patients with various audiometric patterns to average 47 min (range 20–70 min). Estimating an audiogram simultaneously for both ears should not take any longer, unless the thresholds were significantly different between the ears. The 10-nV criterion seems reasonable, but the relationship between criterion noise levels and threshold accuracy needs further study.

## Neurologic applications

The auditory steady-state responses have not been extensively evaluated in patients with neurologic disorders. Several studies have reported that the auditory 40-Hz response is reduced in patients with brainstem or thalamic lesions (Spydell et al, 1985; Rei and Fu, 1988; Harada et al, 1994; Firsching et al, 1987), but not in patients with unilateral lesions of the temporal lobe (Spydell et al, 1985; Rei and Fu, 1988; Firsching et al, 1987). The 40-Hz response is usually absent in comatose patients and in patients with brain death, in keeping with the crucial role of the upper brainstem in generating this response (Firsching et al, 1987; Firsching, 1989). The absence of any effect of temporal lobe lesions does not fit easily with the magnetic and electric source analyses indicating generators in the temporal lobe for the 40-Hz response (see 'Magnetoencephalographic (MEG) recordings', above), or with the animal recordings showing 40-Hz responses in the auditory cortex (Mäkelä et al, 1990; Karmos et al, 1993). The problem can probably be resolved by the existence of cortical generators in both hemispheres. If one generator is removed, the opposite hemisphere may still contribute a reliable response to the vertex recording. More extensive mapping and source analysis would be required to demonstrate the effect of a unilateral lesion. The effects of lesions on responses that reflect the oscillatory interactions between multiple areas may be difficult to understand. The loss of the response with lesions to the brainstem indicates that this drives the oscillations, but the normal response probably involves both cortex and brainstem (see 'Source analysis', above).

Rance et al (1998, 1999) described the steady-state responses in infants and children with auditory neuropathy. These patients had no recognizable brainstem response, but did show cochlear microphonics in response to high-intensity clicks. In these patients, the thresholds for the auditory steady-state responses at rates of 90 Hz were significantly higher than the behavioral thresholds. The physiologic–behavioral differences ranged between –20 and 85 dB, with the greater difference occurring when the behavioral thresholds were lower. Picton et al (1998) described an older child with auditory neuropathy, who had no steady-state responses but near-normal behavioral thresholds.

The associations between neurologic disorders and auditory steady-state responses need to be further evaluated. McAnally and Stein (1997) reported that patients with dyslexia showed smaller steady-state responses than normal subjects for modulation frequencies between 20 and 80 Hz. Borda (1984) found no differences in the 40-Hz response between normal control subjects and patients with degenerative neurologic disorders. However, this might warrant further investigation, since the gamma rhythms of the brain are associated with cognitive processing, and there is some anecdotal evidence that such rhythms are reduced in Alzheimer's disease (Ribary et al, 1991). Investigations of patients with demyelinating disorders may indicate delays in the responses, and the desynchronization associated with demyelination may reduce the amplitude of the responses. Lucertini et al (1993, 2002) showed that the 40-Hz response is delayed by hypobaric hypoxia. Since this response is also affected by anesthesia, it would be worthwhile to evaluate it in patients with metabolic and toxic encephalopathies. In post-traumatic coma, the absence of any 40-Hz response or its disappearance during recovery is a bad prognostic sign (Serafini and Brizi, 1998). The auditory steady-state responses at rapid rates have not been evaluated at all in neurologic patients, other than a few patients with auditory neuropathy.

## Anesthesia applications

The level of anesthesia during surgery is often monitored by means of the responsiveness of the muscular and the cardiovascular systems to stimulation. For many years, anesthetists have searched for some more specific means of monitoring anesthesia that is unaffected by the concomitant neuromuscular blockade or autonomic stabilizers. Unintentional intraoperative awareness occurs when a patient regains consciousness during surgery. This experience can be extremely painful and horrifying. Such episodes may occur without the anesthestist knowing, when voluntary muscle responses are inhibited by neuromuscular blockers and cardiovascular responses are limited by autonomic medication.

Recording the electrical activity of the brain should theoretically provide a more accurate monitoring of the state of anesthesia than reliance upon peripheral responses. However, many of the changes in the EEG with anesthesia are specific to the different types of anesthetic agents. The EEG patterns that reliably indicate the state of anesthesia independent of the anesthetic agents involve combined changes in several different parameters: an increase in delta activity, a decrease in gamma activity, an increase in the coherence between frontal regions, and a decrease in the anterior–posterior and interhemispheric coherence (John et al, 2001c).

Another approach to monitoring anesthesia is to evaluate the response of the brain to sensory stimuli. Auditory stimuli constitute the most reasonable choice for such a procedure, since the stimulation levels can be kept reasonably constant during the operation by using insert earphones. The early auditory responses deriving from the cochlear and brainstem auditory pathways are little affected by anesthetic agents, and are used to monitor the integrity of the auditory pathways during procedures such as the removal of an acoustic neuroma. Since late auditory potentials are variable even during wakefulness, attention has focused on the middle-latency auditory evoked potentials as a possible tool for assessing anesthesia. The middle-latency potentials contain a series of waves or oscillations that recur with a periodicity of 30–40 Hz. The power spectrum of this response when recorded during wakefulness contains most of its energy near 40 Hz. Several studies have looked at these auditory middle-latency responses during anesthesia, and have reported a delay and attenuation of the peaks in the time waveform and a marked decrease in the 40-Hz energy contained in the spectrum (Madler and Pöppel, 1987; Madler et al, 1991; Schwender et al, 1994; Thornton, 1989; Dutton et al, 1999).

These studies, evaluating the transient middle-latency responses evoked by stimuli presented at rates near 10 Hz, were soon supplemented by recordings of the auditory steady-state responses at rates near 40 Hz. Hogan (1987) reported in an abstract that anesthesia with isoflurane and nitrous oxide markedly reduced the auditory 40-Hz steady-state response. Plourde and Picton (1990) showed that the 40-Hz steady-state response was reduced during surgical anesthesia with thiopental, fentanyl, and isoflurane. Furthermore, the amplitude of the response returned as the patients regained consciousness in the recovery room. An underlying hypothesis is that the anesthetic disrupts the natural 40-Hz resonance of the thalamocortical circuits. This fits with the finding that propofol causes dose-dependent changes in regional cerebral bloodflow in both thalamus and sensory cortex (Bonhomme et al, 2001). Since the thalamus acts as a central relay station for afferent sensory information on its way up to the cortex, the disruption of this circuit may serve as a good indicator that the cortex is cut off from aversive sensory stimuli such as pain signals. The steady-state evoked potentials in the 40-Hz range evoked by periodic auditory stimulation may be related to the spontaneous evoked 'gamma'-band activity that occurs in response to a single transient stimulus, which has been associated with conscious perception of an integrated event (Basar-Eroglu et al, 1996). The absence of a recognizable 40-Hz response may be a reliable indicator of unconsciousness, regardless of the exact anesthetic protocol or the concentration of anesthetic agent.

The attenuation of the 40-Hz response during anesthesia was not due to any change in auditory processing in the ear or brainstem, since there were no changes in concomitantly recorded auditory brainstem responses. Nevertheless, it is probably worthwhile to ensure that the auditory input is preserved during the monitoring of the anesthetic, since anesthesia and endotracheal intubation may alter the conduction of sound through the middle ear. One possible way to monitor the auditory input as well as the anesthetic effect on the middle-latency responses would be to record responses to simultaneously presented modulations in the 40-Hz and 80-Hz frequency regions (Lins and Picton, 1995). The 80-Hz response should remain constant, provided that the earphone, middle ear and cochlea are functioning normally, and the 40-Hz response should fluctuate with the level of consciousness.

Since their initial studies using fentanyl and isoflurane, Plourde and co-workers have investigated the auditory steady-state responses during the administration of sufentanil (Gilron et al, 1998), enflurane (Plourde and Villemure, 1996), and propofol (Plourde, 1996). In all of these studies, anesthesia was associated with a profound reduction in the steady-state response. Further studies of the effects of isoflurane in normal volunteers have shown that the attenuation of the steady-state response is related to the end-tidal concentration of isoflurane, and that an absent response or one less than 70 nV in amplitude reliably predicts consciousness (Plourde et al, 1998). Reversing the anesthetic effect of propofol by administering physostigmine caused a return of consciousness in 9 of 11 normal volunteers, and a concomitant increase in the auditory steady-state response (Meuret et al, 2000).

Rather than measuring the auditory steady-state response at a single rate of 40 Hz, Sapsford and co-workers (Sapsford et al, 1996; Andrade et al, 1996) have recorded steady-state responses at multiple stimulus rates and have identified the rate at which the response is the largest. This rate decreases with anesthesia. The natural resonant peak that exists in the 40-Hz range probably disappears with the use of anesthetics, because these substances cause the components of the response to occur at longer latencies. These components then achieve optimum constructive superimposition at lower rates of periodic stimulation.

Interestingly, induction of anesthesia by ketamine has no effect on the transient middle-latency response (Schwender et al, 1994), and actually increases the amplitude of the 40-Hz auditory steady-state response (Plourde et al, 1997; Conti et al, 1998). These results show clearly that unconsciousness can occur in the presence of normal auditory steady-state responses. The logic for using these responses to monitor anesthesia is therefore not so much that the responses indicate consciousness, but that the absence of the responses (provided that the auditory input is normal) indicates unconsciousness. The results reinforce the idea that ketamine, which is a dissociative anesthetic, induces unconsciousness in a manner quite different from other anesthetic agents.

The 40-Hz steady-state response is probably one example of the 30–40-Hz rhythmic activities that are set up in thalamo-cortical systems (Kulli and Koch, 1991; Plourde, 1999). These rhythmic oscillations are characteristic of the activity in thalamocortical neuronal networks (Steriade et al, 1998), and are increased during brain activation (Steriade et al, 1996) and reduced by anesthesia (Tennigkeit et al, 1997).

## Overview

Auditory steady-state responses can be recorded from the human scalp at stimulus rates up to several hundred hertz. Since these responses are exactly locked to the stimulus rate, they are best detected after the recording has been transferred to the frequency domain. The most prominent responses occur at rates near 40 Hz. Responses at faster rates may also be easily recognized, even though they are smaller, since the background

EEG noise becomes lower at higher frequencies. The 40-Hz responses are probably generated by interactions between the thalamus and auditory cortices, whereas the main generators of the faster responses are probably located in the brainstem. The 40-Hz response is decreased by sleep and anesthesia, whereas the 80-Hz response remains stable.

The auditory steady-state responses have demonstrated usefulness in two areas. First, they can provide objective demonstration that sounds have been processed by the brain, and thus contribute to objective audiometry. Here, the responses at higher stimulus rates are more useful, since they are not affected by sleep and are easily recorded in newborn infants. However, like the auditory brainstem responses, the auditory steady-state responses may not be able to estimate hearing thresholds correctly in patients with auditory neuropathy. Second, the 40-Hz responses are useful in demonstrating that the cortex is responding normally to sounds. This is helpful in audiometry and important in detecting inadvertent intraoperative awareness during anesthesia.

Much still needs to be determined before we can fully understand these responses and use them more extensively. Further developments in signal analysis should lead to more efficient techniques for detecting these responses. Evoking these responses using different stimuli (e.g. involving binaural processing) may show the physiologic bases for different types of auditory perception. A deeper understanding of the intracerebral generators of the responses might provide a basis for using the responses to localize abnormalities in the central auditory nervous system.

Probably the most pressing need at the present time is for more extensive normative data for audiometric applications. Most of the threshold estimations in the literature are from adults or older children. We need many more normative data for newborn babies and young infants, so that we may more accurately assess hearing and development in these subjects. This is especially true for infants with hearing loss. We also need to study the bone conduction responses more intensively, in relation to thresholds and artefacts. Which stimulus rates might be more appropriate at different ages and for different reasons need to be determined. Is 40 Hz the better rate in adults, and rates of greater than 80 Hz better in infants? The technique is ready to be used for clinical audiometry, but it should be used cautiously. The procedures can still be improved by more experimental work. There remains much to be done.

## Acknowledgments

Barbara Cone-Wesson, Robert Dobie, Anthony Norcia, Michael Scherg, Christo Pantev, Bernhard Roß, Marilyn Perez-Abalo, and Pedro Valdes-Sosa.

## References

Anderson TW (1984). An introduction to multivariate statistical analysis. New York: Wiley.

Andrade J, Sapsford DJ, Jeevaratnum D, Pickworth AJ, Jones JG (1996). The coherent frequency in the electroencephalogram as an objective measure of cognitive function during propofol sedation. Anesth Analg 83:1279–1284.

Aoyagi M, Fuse T, Suzuki T, Kim Y, Koike Y (1993a). An application of phase spectral analysis to amplitude-modulation following response. Acta Otolaryngol Suppl 504:82–88.

Aoyagi M, Kiren T, Kim Y, Suzuki Y, Fuse T, Koike Y (1993b). Optimal modulation frequency for amplitude-modulation following response in young children during sleep. Hear Res 65:253–261.

Aoyagi M, Kiren T, Kim Y, Suzuki Y, Fuse T, Koike Y (1993c). Frequency specificity of amplitude-modulation-following response detected by phase spectral analysis. Audiology 32:293–301.

Aoyagi M, Kiren T, Furuse H, et al (1994a). Effects of aging on amplitude-modulation following response. Acta Otolaryngol Suppl 511:15–22.

Aoyagi M, Furuse H, Yokota M, Kiren T, Suzuki Y, Koike Y (1994b). Detectability of amplitude-modulation following response at different carrier frequencies. Acta Otolaryngol Suppl 511:23–27.

Aoyagi M, Kiren T, Furuse H, et al (1994c). Pure-tone threshold prediction by 80-Hz amplitude-modulation following response. Acta Otolaryngol Suppl 511:7–14.

Aoyagi M, Suzuki Y, Yokota M, Furuse H, Watanabe T, Ito T (1999). Reliability of 80-Hz amplitude-modulation-following response detected by phase coherence. Audiol Neurootol 4:28–37.

Arnold S, Burkard R (1998). The auditory evoked potential difference tone and cubic difference tone measured from the inferior colliculus of the chincilla. J Acoust Soc Am 104:1565–1573.

Azzena GB, 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.

Barajas JJ, Fernandez R, Bernal MR (1988). Middle latency and 40 Hz auditory evoked responses in normal hearing children: 500 Hz thresholds. Scand Audiol Suppl 30:99–104.

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.

Basar-Eroglu C, Struber D, Schurmann M, Stadler M, Basar E (1996). Gamma-band responses in the brain: a short review of psychophysiological correlates and functional significance. Int J Psychophysiol 24:101–112.

Batra R, Kuwada S, Maher VL (1986). The frequency-following response to continuous tones in humans. Hear Res 21:167–177.

Bijl GK, Veringa F (1985). Neural conduction time and steady-state evoked potentials. Electroencephalogr Clin Neurophysiol 62:465–467.

Boettcher FA, Poth EA, Mills JH, Dubno JR (2001). The amplitude-modulation following response in young and aged human subjects. Hear Res 153:32–42.

Bonhomme V, Fiset P, Meuret P, et al (2001). Propofol anesthesia and cerebral blood flow changes elicited by vibrotactile stimulation: a positron emission tomography study. J Neurophysiol 85: 1299–1308.

Borda RP (1984). The 40/sec middle-latency auditory response in Alzheimer's disease, Parkinson's disease, and age-matched normals. PhD thesis, University of Houston, Houston.

Brillinger DR (1975). Time series: data analysis and theory. New York: Holt, Rinehart and Winston.

Campbell FW, Atkinson J, Francis MR, Green DM (1977). Estimation of auditory thresholds using evoked potentials. A clinical screening test. Prog Clin Neurophysiol 2:68–78.

Cebulla M, Stürzebecher E, Wernecke KD (2001). Objective detection

of amplitude-modulation following response (AMFR). Audiology 40:245–252.

Chambers RD, Meyer TA (1993). Reliability of threshold estimation in hearing-impaired adults using the AMFR. J Am Acad Audiol 4:22–32.

Champlin CA (1992). Method for detecting auditory steady-state potentials recorded from humans. Hear Res 58:63–69.

Chatrian GE, Petersen MC, Lazarte JA (1960). Responses to clicks from the human brain: some depth electrographic observations. Electroencephalogr Clin Neurophysiol 12:479–489.

Chertoff ME, Hecox KE (1990). Auditory nonlinearities measured with auditory-evoked potentials. J Acoust Soc Am 87:1248–1254.

Chertoff ME, Hecox KE, Goldstein R (1992). Auditory distortion products measured with averaged auditory evoked potentials. J Speech Hear Res 35:157–166.

Ciulla C, Takeda T, Morabito M, Endo H, Kumagai T, Xiao R (1996). MEG measurements of 40 Hz auditory evoked response in human brain. Electroencephalogr Clin Neurophysiol Suppl 47:121–127.

Cohen LT, Rickards FW, Clark GM (1991). A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. J Acoust Soc Am 90:2467–2479.

Cone-Wesson B, Dowell RC, Tomlin D, Rance G, Ming WJ (2002a). The auditory steady-state response: I Comparisons with the auditory brainstem response. J Am Acad Audiol 13:173–187.

Cone-Wesson B, Parker J, Swiderski N, Rickards F (2002b). The auditory steady-state response: II. Full-term and premature neonates. J Am Acad Audiol 13:260–269.

Cone-Wesson B, Rickards F, Poulis C, Parker J, Tan L, Pollard J (2002c). The auditory steady-state response: III Clinical observations and applications in infants and children. J Am Acad Audiol 13:270–282.

Conti G, Santarelli R, Ottaviani F, Azzena GB (1998). Generation of human auditory steady-state responses. In: Litscher G, ed. Monitoring of stimulus-induced 40Hz brain oscillations. Scottsdale: Pabst Science Publishers, pp. 84–107.

Conti G, Santarelli R, Grassi C, Ottaviani F, Azzena GB (1999). Auditory steady-state responses to click trains from the rat temporal cortex. Clin Neurophysiol 110:62–70.

Cooley JW, Tukey JW (1965). An algorithm for the machine calculation of complex Fourier series. Math Comput 19:297–301.

Dauman R, Szyfter W, Charlet de Sauvage R, Cazals Y (1984). Low frequency thresholds assessed with 40 Hz MLR in adults with impaired hearing. Arch Otorhinolaryngol 240:85–89.

Diamond AL (1977). Latency of the steady state visual evoked potential. Electroencephalogr Clin Neurophysiol 42:125–127.

Dimitrijevic A, John MS, van Roon P, Picton TW (2001). Human auditory steady-state responses to tones independently modulated in both frequency and amplitude. Ear Hear 22:100–111.

Dimitrijevic A, John MS, Van Roon P, et al (2002). Estimating the audiogram using multiple auditory steady-state responses. J Am Acad Audiol 13:205–224.

Dobie RA, Wilson MJ (1988). Auditory responses to the envelopes of pseudorandom noise stimuli in humans. Hear Res 36:9–20.

Dobie RA, Wilson MJ (1989). Analysis of auditory evoked potentials by magnitude-squared coherence. Ear Hear 10:2–13.

Dobie RA, Wilson MJ (1991). Optimal smoothing of coherence estimates. Electroencephalogr Clin Neurophysiol 80:194–200.

Dobie RA, Wilson MJ (1993). Objective response detection in the frequency domain. Electroencephalogr Clin Neurophysiol 88:516–524.

Dobie RA, Wilson MJ (1994a). Objective detection of 40 Hz auditory evoked potentials: phase coherence vs. magnitude-squared coherence. Electroencephalogr Clin Neurophysiol 92:405–413.

Dobie RA, Wilson MJ (1994b). Phase weighting: a method to improve objective detection of steady-state evoked potentials. Hear Res 79:94–98.

Dobie RA, Wilson MJ (1995a). Objective versus human observer detection of 40-Hz auditory-evoked potentials. J Acoust Soc Am 97:3042–3050.

Dobie RA, Wilson MJ (1995b). Comparison of objective threshold estimation procedures for 40-Hz auditory evoked potentials. Ear Hear 16:299–310.

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

Dolphin WF (1996). Auditory evoked responses to amplitude modulated stimuli consisting of multiple envelope components. J Comp Physiol A 179:113–121.

Dolphin WF (1997). The envelope following response to multiple tone pair stimuli. Hear Res 110:1–14.

Dolphin WF, Mountain DC (1993). The envelope following response (EFR) in the Mongolian gerbil to sinusoidally amplitude-modulated signals in the presence of simultaneously gated pure tones. J Acoust Soc Am 94:3215–3226.

Dolphin WF, Chertoff ME, Burkard R (1994). Comparison of the envelope following response in the Mongolian gerbil using two-tone and sinusoidally amplitude-modulated tones. J Acoust Soc Am 96:2225–2234.

Dolphin WF, Au WWL, Nachtigall PE, Pawloski J (1995). Modulation rate transfer functions to low-frequency carriers in three species of cetaceans. J Comp Physiol A 177:235–245.

Dumermuth G, Molinari L (1987). Spectral analysis of EEG background activity. In: Gevins AS, Remond A, eds. Methods of analysis of brain electrical and magnetic signals. Amsterdam: Elsevier, pp. 85–130.

Dutton RC, Smith WD, Rampil IJ, Chortkoff BS, Eger EI (1999). Forty-Hertz midlatency auditory evoked potential activity predicts wakeful response during desflurane and propofol anesthesia in volunteers. Anesthesiology 91:1209–1220.

Eggermont JJ (2001). Between sound and perception: reviewing the search for a neural code. Hear Res 157:1–42.

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

Fastl H (1990). The hearing sensation roughness and neuronal responses to AM-tones. Hear Res 46:293–295.

Firsching R (1989). The brain-stem and 40 Hz middle latency auditory evoked potentials in brain death. Acta Neurochir 101:52–55.

Firsching R, Luther J, Eidelberg E, Brown WE Jr, Story JL, Boop FA (1987). 40 Hz-middle latency auditory evoked response in comatose patients. Electroencephalogr Clin Neurophysiol 67:213–216.

Fisher RA (1929). Tests of significance in harmonic analysis. Proc R Soc Lond 125:54–59.

Fisher NI (1993). Statistical analysis of circular data. Cambridge: Cambridge University Press.

Franowicz MN, Barth DS (1995). Comparison of evoked potentials and high-frequency (gamma-band) oscillating potentials in rat auditory cortex. J Neurophysiol 74:96–112.

Fridman J, Zappulla R, Bergelson M, et al (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 (1981). A 40-Hz auditory potential recorded from the human scalp. Nicolet Potentials 1:12.

Galambos R (1992). A comparison of certain gamma band (40-Hz) brain rhythms in cat and man. In: Basar E, Bullock TH, eds. Induced rhythms in the brain. Boston: Birkhäuser, pp. 201–216.

Galambos R, Makeig S (1988). Dynamic changes in steady-state potentials. In: Basar E, ed. Dynamics of sensory and cognitive processing of the brain. Berlin: Springer, pp. 178–199.

Galambos R, Makeig S (1992a). Physiological studies of central masking in man. I: The effects of noise on the 40-Hz steady-state response. J Acoust Soc Am 92:2683–2690.

Galambos R, Makeig S (1992b). Physiological studies of central masking in man. II: Tonepip SSRS and the masking level difference. J Acoust Soc Am 92:2691–2697.

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

Geisler CD (1960). Average responses to clicks in man recorded by

scalp electrodes. Technical Report 380. Cambridge: MIT Research Laboratories of Electronics.

Gilron I, Plourde G, Marcantoni W, Varin F (1998). 40 Hz auditory steady-state response and EEG spectral edge frequency during sufentanil anaesthesia. Can J Anaesth 45:115–121.

Giraud AL, Lorenzi C, Ashburner J, et al (2000). Representation of the temporal envelope of sounds in the human brain. J Neurophysiol 84:1588–1598.

Goldstein JL, Baer T, Kiang NYS (1971). A theoretical treatment of latency, group delay, and tuning characteristics for auditory-nerve responses to clicks and tones. In: Sachs MB, ed. Physiology of the auditory system; based on the proceedings of a workshop. Baltimore: National Educational Consultants, pp. 133–141.

Gorga MP, Thornton AR (1989). The choice of stimuli for ABR measurements. Ear Hear 10:217–230.

Gorga MP, Beauchaine KA, Reiland JK (1987). Comparison of onset and steady-state responses of hearing aids: implications for the use of the auditory brainstem response in the selection of hearing aids. J Speech Hear 30:130–136.

Green GG, Kay RH, Rees A (1979). Responses evoked by frequency-modulated sounds recorded from the human scalp. J Physiol 296: 21P–22P.

Griffiths SK, Chambers RD (1991). The amplitude modulation-following response as an audiometric tool. Ear Hear 12:235–241.

Gutschalk A, Mase R, Roth R, et al (1994). Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clin Neurophysiol 110:856–868.

Hall JW (1979). Auditory brainstem frequency following responses to waveform envelope periodicity. Science 205:1297–1299.

Harada J, Aoyagi M, Suzuki T, Kiren T, Koike Y (1994). A study on the phase spectral analysis of middle latency response and 40-Hz event-related potential in central nervous system disorders. Acta Otolaryngol Suppl 511:34–39.

Hari R, Hämäläinen M, Joutsiniemi SL (1989). Neuromagnetic steady-state responses to auditory stimuli. J Acoust Soc Am 86:1033–1039.

Herdman AT, Stapells DR (2001). Thresholds determined using the monotic and dichotic multiple auditory steady-state response technique in normal-hearing subjects. Scand Audiol 30:41–49.

Herdman AT, Lins O, Van Roon P, Stapells DR, Scherg M, Picton TW (2002a). Intracerebral sources of human auditory steady-state responses. Brain Topography 15:69–86.

Herdman AT, Picton TW, Stapells DR (2002b). Place specificity of auditory steady state responses. J Acoust Soc Am 112:1569–1582.

Herdman AT, Stapells DR (2003). Auditory steady state response thresholds of adults with sensorineural hearing impairment. Int J Audiol (in press).

Hink RF, Kodera K, Yamada O, Kaga K, Suzuki J (1980). Binaural interaction of a beating frequency-following response. Audiology 19:36–43.

Hogan K (1987). 40 Hz steady-state evoked potentials (SSEP) during isoflurane-N2O anesthesia. Anesthesiology 67:A402.

Hoke M, Ross B, Wickesberg R, Lütkenhöner B (1984). Weighted averaging—theory and application to electric response audiometry. Electroencephalogr Clin Neurophysiol 57:484–489.

Hotelling H (1931). The generalization of student's ratio. Ann Math Statist 2:360–378.

Isaksson A, Wennberg A (1975). Visual evaluation and computer analysis of the EEG—a comparison. Electroencephalogr Clin Neurophysiol 38:79–86.

Jerger J, Chmiel R, Frost JD Jr, Coker N (1986). Effect of sleep on the auditory steady state evoked potential. Ear Hear 7:240–245.

Jerger J, Chmiel R, Glaze D, Frost JD Jr (1987). Rate and filter dependence of the middle-latency response in infants. Audiology 26:269–283.

John MS, Picton TW (2000a). Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements. Hear Res 141:57–79.

John MS, Picton TW (2000b). MASTER: a windows program for recording multiple auditory steady-state responses. Comput Methods Programs Biomed 61:125–150.

John MS, Lins OG, Boucher BL, Picton TW (1998). Multiple auditory steady-state responses (MASTER): stimulus and recording parameters. Audiology 37:59–82.

John MS, Dimitrijevic A, Picton TW (2001a). Weighted averaging of steady-state responses. Clin Neurophysiol 112:555–562.

John MS, Dimitrijevic A, van Roon P, Picton TW (2001b). Multiple auditory steady-state responses to AM and FM stimuli. Audiol Neurootol 6:12–27.

John ER, Prichep LS, Kox W, et al (2001c). Invariant reversible qEEG effects of anesthetics. Consciousness Cognition 10:165–183.

John MS, Dimitrijevic A, Picton TW (2002a). Auditory steady-state responses to exponential modulation envelopes. Ear Hear 23:106–117.

John MS, Purcell DW, Dimitrijevic A, Picton TW (2002b). Advantages and caveats when recording steady-state responses to multiple simultaneous stimuli. J Am Acad Audiol 13:246–259.

Johnson BW, Weinberg H, Ribary U, Cheyne DO, Ancill R (1988). Topographic distribution of the 40 Hz auditory evoked-related potential in normal and aged subjects. Brain Topogr 1:117–121.

Kankkunen A, Rosenhall U (1985). Comparison between thresholds obtained with pure-tone audiometry and the 40-Hz middle latency response. Scand Audiol 14:99–104.

Karmos G, Mäkelä JP, Ulbert I, Winkler I (1993). Evidence for intracortical generation of the auditory 40-Hz response in the cat. In: Heinze HJ, Münte TF, Mangun GR, eds. New developments in event-related potentials. Boston: Birkhäuser, pp. 87–93.

Khanna SM, Teich MC (1989a). Spectral characteristics of the responses of primary auditory-nerve fibers to amplitude-modulated signals. Hear Res 39:143–157.

Khanna SM, Teich MC (1989b). Spectral characteristics of the responses of primary auditory-nerve fibers to frequency-modulated signals. Hear Res 39:159–175.

Kileny P, Shea SL (1986). Middle-latency and 40-Hz auditory evoked responses in normal-hearing subjects: click and 500-Hz thresholds. J Speech Hear Res 29:20–28.

Kiren T, Aoyagi M, Furuse H, Koike Y (1994). An experimental study on the generator of amplitude-modulation following response. Acta Otolaryngol Suppl 511:28–33.

Klein AJ (1983a). Properties of the brain-stem response slow-wave component. I. Latency, amplitude and threshold sensitivity. Arch Otolaryngol 109:6–12.

Klein AJ (1983b). Properties of the brain-stem response slow-wave component. II. Frequency specificity. Arch Otolaryngol 109:74–78.

Knight RD, Kemp DT (2000). Indications of different distortion product otoacoustic emission mechanisms from a detailed f1, f2 area study. J Acoust Soc Am 107:457–473.

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

Krishnan A, Parkinson J (2000). Human frequency-following response: representation of tonal sweeps. Audiol Neurootol 5:312–321.

Kulli J, Koch C (1991). Does anesthesia cause loss of consciousness? Trends Neurosci 14:6–10.

Kuwada S, Batra R, Maher VL (1986). Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones. Hear Res 21:179–192.

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

Lehmann D, Michel CM (1989). Intracerebral dipole sources of EEG FFT power maps. Brain Topogr 2:155–164.

Lehmann D, Michel CM (1990). Intracerebral dipole source localization for FFT power maps. Electroencephalogr Clin Neurophysiol 76: 271–276.

Lenarz T, Gulzow J, Grozinger M, Hoth S (1986). Clinical evaluation of 40-Hz middle-latency responses in adults: frequency specific threshold estimation and suprathreshold amplitude characteristics, Otorhinolaryngol Relat Spec ORL 48:24–32.

Levi EC, Folsom RC, Dobie RA (1993). Amplitude-modulation following response (AMFR): effects of modulation rate, carrier frequency, age and state. Hear Res 68:42–52.

Levi EC, Folsom RC, Dobie RA (1995). Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. Hear Res 89:21–27.

Linden RD, Campbell KB, Hamel G, Picton TW (1985). Human auditory steady state evoked potentials during sleep. Ear Hear 6:167–174.

Linden RD, Picton TW, Hamel G, Campbell KB (1987). Human auditory steady state evoked potentials during selective attention. Electroencephalogr Clin Neurophysiol 66:145–159.

Lins OG, Picton TW (1995). Auditory steady-state responses to multiple simultaneous stimuli. Electroencephalogr Clin Neurophysiol 96:420–432.

Lins OG, Picton PE, Picton TW, Champagne SC, Durieux-Smith A (1995). Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz. J Acoust Soc Am 97:3051–3063.

Lins OG, Picton TW, Boucher BL, et al (1996). Frequency-specific audiometry using steady-state responses. Ear Hear 17:81–96.

Lucertini M, Ciniglio Appiani G, Antonini R, Urbani L (1993). Effects of hypobaric hypoxia on the middle-latency and steady-state auditory evoked potentials. Audiology 32:356–362.

Lucertini M, Verde P, De Santis S (2002). Human auditory steady-state responses during repeated exposure to hypobaric hypoxia. Audiol Neurootol 7:107–113.

Lütkenhöner B (1992). Frequency-domain localization of intracerebral dipolar sources. Electroencephalogr Clin Neurophysiol 82:112–118.

Lütkenhöner B, Hoke M, Pantev C (1985). Possibilities and limitations of weighted averaging. Biol Cybern 52:409–416.

Lynn JM, Lesner SA, Sandridge SA, Daddario CC (1984). Threshold prediction from the auditory 40-Hz evoked potential. Ear Hear 5:366–370.

Madler C, Pöppel E (1987). Auditory evoked potentials indicate the loss of neuronal oscillations during general anaesthesia. Naturwissenschaften 74:42–43.

Madler C, Keller I, Schwender D, Pöppel E (1991). Sensory information processing during general anaesthesia: effect of isoflurane on auditory evoked neuronal oscillations. Br J Anaesth 66:81–87.

Maiste A, Picton T (1989). Human auditory evoked potentials to frequency-modulated tones. Ear Hear 10:153–160.

Makeig S, Galambos R (1989). The CERP: event-related perturbations in steady-state responses. In: Basar E, ed. Brain dynamics: progress and perspectives. Berlin: Springer, pp. 373–400.

Makeig S, Müller MM, Rockstroh B (1996). Effects of voluntary movements on early auditory brain responses. Exp Brain Res 110:487–492.

Mäkelä JP, Hari R (1987). Evidence for cortical origin of the 40 Hz auditory evoked response in man. Electroencephalogr Clin Neurophysiol 66:539–546.

Mäkelä JP, Karmos G, Molnar M, Csepe V, Winkler I (1990). Steady-state responses from the cat auditory cortex. Hear Res 45:41–50.

Mardia KV (1972). Statistics of directional data. New York: Academic Press.

Mardia KV, Kent JT, Bibby JM (1979). Multivariate analysis. New York: Academic Press.

Mauer G, Döring WH (1999). Generators of amplitude modulation following response (AMFR). Paper presented at 16th meeting of the Evoked Response Audiometry Study Group, Tromsø, Norway, June, 1999.

Maurizi M, Almadori G, Paludetti G, Ottaviani F, Rosignoli M, Luciano R (1990). 40-Hz steady-state responses in newborns and in children. Audiology 29:322–328.

McAnally KI, Stein JF (1997). Scalp potentials evoked by amplitude modulated tones in dyslexia. J Speech Lang Hear Res 40:939–945.

McGee T, Kraus N (1996). Auditory development reflected by middle latency response. Ear Hear 17:419–429.

McGee T, Kraus N, Killion M, Rosenberg R, King C (1993). Improving the reliability of the auditory middle latency response by monitoring EEG delta activity. Ear Hear 14:76–84.

Meuret P, Backman SB, Bonhomme V, Plourde G, Fiset P (2000). Physostigmine reverses propofol-induced unconsciousness and attenuation of the auditory steady state response and bispectral index in human volunteers. Anesthesiology 93:708–717.

Milford CA, Birchall JP (1989). Steady-state auditory evoked potentials to amplitude-modulated tones in hearing-impaired subjects. Br J Audiol 23:137–142.

Moore BC, Glasberg BR (1988). Gap detection with sinusoids and noise in normal, impaired, and electrically stimulated ears. J Acoust Soc Am 83:1093–1101.

Moushegian G, Rupert AL, Stillman RD (1973). Scalp-recorded early responses in man to frequencies in the speech range. Electroencephalogr Clin Neurophysiol 35:665–667.

Muchnik C, Katz-Putter H, Rubinstein M, Hildesheimer M (1993). Normative data for 40-Hz event-related potentials to 500-Hz tonal stimuli in young and elderly subjects. Audiology 32:27–35.

Munglani R, Andrade J, Sapsford DJ, Baddeley A, Jones JG (1993). A measure of consciousness and memory during isoflurane administration: the coherent frequency. Br J Anaesth 71:633–641.

Norcia AM, Tyler CW, Hamer RD, Wesemann W (1989). Measurement of spatial contrast sensitivity with the swept contrast VEP. Vision Res 29:627–637.

Oates P, Stapells DR (1997). Frequency specificity of the human auditory brainstem and middle latency responses to brief tones. II. Derived response analyses. J Acoust Soc Am 102:3609–3619.

Palaskas CW, Wilson MJ, Dobie RA (1989). Electrophysiologic assessment of low-frequency hearing: sedation effects. Otolaryngol Head Neck Surg 101:434–441.

Pantev C, Elbert T, Makeig S, Hampson S, Eulitz C, Hoke M (1993). Relationship of transient and steady-state auditory evoked fields. Electroencephalogr Clin Neurophysiol 88:389–396.

Pantev C, Roberts LE, Elbert T, Ross B, Wienbruch C (1996). Tonotopic organization of the sources of human auditory steady-state responses. Hear Res 101:62–74.

Parker D, Matsebula D (1998). The period evoked potential: objective response detection and 500 Hz thresholds in normally hearing adults. Br J Audiol 32:167–175.

Parker D, O'Dwyer D (1998). The 40 Hz modulation-following response: prediction of low-frequency uncomfortable loudness levels in normally hearing adults. Audiology 37:372–381.

Perez-Abalo MC, Savio G, Torres A, Martin V, Rodriguéz E, Galan L (2001). Steady state responses to multiple amplitude-modulated tones: an optimized method to test frequency-specific thresholds in hearing-impaired children and normal-hearing subjects. Ear Hear 22:200–211.

Pethe J, Hocke T, Müller R, von Specht H (2000). On the frequency spectrum of amplitude modulation following responses. Scand Audiol 29:191–195.

Pethe J, Von Specht H, Mühler R, Hocke T (2001). Amplitude modulation following responses in awake and sleeping humans—a comparison for 40 Hz and 80 Hz modulation frequency. Scand Audiol Suppl 52:152–155.

Picton TW (1978). The strategy of evoked potential audiometry. In: Gerber SE, Mencher GT, eds. Early diagnosis of hearing loss. New York: Grune and Stratton, pp. 297–307.

Picton TW (1987). Human auditory steady state responses. In: Barber C, Blum T, eds. Evoked potentials III. Boston: Butterworth, pp. 117–124.

Picton TW, Rodriguez RT, Linden RD, Maiste AC (1985). The neurophysiology of human hearing. Human Communication Canada 9:127–136.

Picton TW, Vajsar J, Rodriguez R, Campbell KB (1987a). Reliability estimates for steady state evoked potentials. Electroencephalogr Clin Neurophysiol 68:119–131.

Picton TW, Dauman R, Aran J-M (1987b). Réponses évoquées en 'régime permanent' chez l'homme par la modulation sinusoïdale de fréquence. J Otolaryngol 16:140–145.

Picton TW, Skinner CR, Champagne SC, Kellett AJ, Maiste AC (1987c). Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. J Acoust Soc Am 82:165–178.

Picton TW, Durieux-Smith A, Champagne SC, et al (1998). Objective evaluation of aided thresholds using auditory steady-state responses. J Am Acad Audiol 9:315–331.

Picton TW, Dimitrijevic A, John MS, van Roon P (2001). The use of phase in the detection of auditory steady-state responses. Clin Neurophysiol 112:1692–1711.

Picton TW, Dimitrijevic A, John MS (2002a). Multiple auditory steady-state responses. Ann Otol Rhinol Laryngol 111(suppl 189):16–21.

Picton TW, Dimitrijevic A, Van Roon P, John MS, Reed M, Finklestein H (2002b). Possible roles for the auditory steady-state responses in fitting hearing aids. In: Seewald RC, Gravel JS, eds. A sound foundation through early amplification 2001. Proceedings of the 2nd international conference. Basel: Phonak AG, pp. 63–73.

Plourde G (1996). The effects of propofol on the 40-Hz auditory steady-state response and on the electroencephalogram in humans. Anesth Analg 82:1015–1022.

Plourde G (1999). Auditory evoked potentials and 40-Hz oscillations. Anesthesiology 91:1187–1189.

Plourde G, Picton TW (1990). Human auditory steady state responses during general anesthesia. Anesth Analg 71:460–468.

Plourde G, Villemure C (1996). Comparison of the effects of enflurane/N2O on the 40-Hz auditory steady-state response versus the auditory middle-latency response. Anesth Analg 82:75–83.

Plourde G, Stapells DR, Picton TW (1991). The human auditory steady-state evoked potentials. Acta Otolaryngol Suppl 491:153–159.

Plourde G, Baribeau J, Bonhomme V (1997). Ketamine increases the amplitude of the 40-Hz auditory steady-state response in humans. Br J Anaesth 78:524–529.

Plourde G, Villemure C, Fiset P, Bonhomme V, Backman SB (1998). Effect of isoflurane on the auditory steady-state response and on consciousness in human volunteers. Anesthesiology 89:844–851.

Poon PW, Chiu TW (2000). Similarities of FM and AM receptive space of single units at the auditory midbrain. Biosystems 58:229–237.

Popov VV, Supin AY, Klishin VO (1996). Frequency tuning curves of the dolphin's hearing: envelope-following response study. J Comp Physiol [A] 178:571–578.

Purcell DW, John MS, Picton TW (2003). Concurrent measurement of distortion product otoacoustic emissions and auditory steady state evoked potentials. Hear Res 176:128–141.

Rance G, Briggs RJS (2002). Assessment of hearing in infants with moderate to profound impairment: the Melbourne experience with auditory steady-state evoked potential testing. Ann Otol Rhinol Laryngol 111(suppl 189):22–28.

Rance G, Rickards F (2002). Prediction of hearing threshold in infants using auditory steady-state evoked potentials. J Am Acad Audiol 13:236–245.

Rance G, Rickards FW, Cohen LT, Burton MJ, Clark GM (1993). Steady state evoked potentials: a new tool for the accurate assessment of hearing in cochlear implant candidates. Adv Otorhinolaryngol 48:44–48.

Rance G, Rickards FW, Cohen LT, De Vidi S, Clark GM (1995). The automated prediction of hearing thresholds in sleeping subjects using auditory steady-state evoked potentials. Ear Hear 16:499–507.

Rance G, Dowell RC, Rickards FW, Beer DE, Clark GM (1998). Steady-state evoked potential and behavioral hearing thresholds in a group of children with absent click-evoked auditory brain stem response. Ear Hear 19:48–61.

Rance G, Beer DE, Cone-Wesson B, et al (1999). Clinical findings for a group of infants and young children with auditory neuropathy. Ear Hear 20:238–252.

Rayleigh L (1880). On the resultant of a large number of vibrations of the same pitch and arbitrary phase. Phil Mag 10:73–78.

Rees A, Green GG, Kay RH (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. Hear Res 23:123–133.

Regan D (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. Electroencephalogr Clin Neurophysiol 20:238–248.

Regan D (1973). Rapid objective refraction using evoked brain potentials. Invest Opthalmol 12:669–679.

Regan D (1976). Latencies of evoked potentials to flicker and to pattern speedily estimated by simultaneous method. Electroencephalogr Clin Neurophysiol 40:654–660.

Regan D (1977). Evoked potentials in basic and clinical research. In: Rémond A, ed. EEG informatics: a didactic review of methods and applications of EEG data processing. Amsterdam: Elsevier, pp. 319–346.

Regan D (1989). Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine. Amsterdam: Elsevier.

Regan MP (1994a). Linear half-wave rectification of modulated sinusoids. Appl Math Comput 62:61–79.

Regan MP (1994b). A method for calculating the spectral response of a hair cell to a pure tone. Biol Cybern 71:13–16.

Regan D, Heron JR (1969). Clinical investigation of lesions of the visual pathway: a new objective technique. J Neurol Neurosurg Psychiatry 32:479–483.

Regan MP, Regan D (1988). A frequency domain technique for characterizing nonlinearities in biological systems. J Theor Biol 133:293–317.

Regan MP, Regan D (1993). Nonlinear terms produced by passing amplitude-modulated sinusoids through a hair cell transducer function. Biol Cybern 69:439–446.

Rei G, Fu BT (1988). Diagnostic significance of the staggered spondaic word test and 40-Hz auditory event-related potentials. Audiology 27:8–16.

Ribary U, Ioannides AA, Singh KD, et al (1991). Magnetic field tomography of coherent thalamocortical 40-Hz oscillations in humans. Proc Natl Acad Sci USA 88:11037–11041.

Rickards FW, Clark GM (1984). Steady-state evoked potentials to amplitude-modulated tones. In: Nodar RH, Barber C, eds. Evoked potentials II. Boston, MA: Butterworth, pp. 163–168.

Rickards FW, Tan LE, Cohen LT, Wilson OJ, Drew JH, Clark GM (1994). Auditory steady-state evoked potential in newborns. Br J Audiol 28:327–337.

Rockstroh B, Muller M, Heinz A, Wagner M, Berg P, Elbert T (1996). Modulation of auditory responses during oddball tasks. Biol Psychol 43:41–55.

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.

Rohrbaugh JW, Varner JL, Paige SR, Eckardt MJ, Ellingson RJ (1989). Event-related perturbations in an electrophysiological measure of auditory function: a measure of sensitivity during orienting? Biol Psychol 29:247–271.

Rohrbaugh JW, Varner JL, Paige SR, Eckardt MJ, Ellingson RJ (1990). Auditory and visual event-related perturbations in the 40 Hz auditory steady-state response. Electroencephalogr Clin Neurophysiol 76:148–164.

Roß B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000). A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J Acoust Soc Am 108:679–691.

Roß B, Picton TW, Pantev C (2002). The onset of steady-state responses reflects temporal integration in the human auditory cortex. Hear Res 165:68–84.

Sammeth CA, Barry SJ (1985). The 40-Hz event-related potential as a measure of auditory sensitivity in normals. Scand Audiol 14:51–55.

Santarelli R, Conti G (1999). Generation of auditory steady-state responses: linearity assessment. Scand Audiol Suppl 51:23–32.

Santarelli R, Maurizi M, Conti G, Ottaviani F, Paludetti G, Pettorossi VE (1995). Generation of human auditory steady-state responses (SSRS). II: Addition of responses to individual stimuli. Hear Res 83:9–18.

Sapsford DJ, Pickworth AJ, Jones JG (1996). A method for producing the coherent frequency: a steady-state auditory evoked response in the electroencephalogram. Anesth Analg 83:1273–1278.

Savio G, Cardenas J, Perez Abalo M, Gonzalez A, Valdes J (2001). The low and high frequency auditory steady state responses mature at different rates. Audiol Neurootol 6:279–287.

Sayers BM, Beagley HA, Riha J (1979). Pattern analysis of auditory-evoked EEG potentials. Audiology 18:1–16.

Schimmel H, Rapin I, Cohen MM (1974). Improving evoked response audiometry with special reference to the use of machine scoring. Audiology 13:33–65.

Schoonhoven R, Boden CJR, Verbunt JPA, de Munck JC (2001). A whole-head MEG study of the amplitude modulation following response: phase coherence, group delays and dipole source analysis. Paper presented at 17th meeting of the Evoked Response Audiometry Study Group, Vancouver, Canada, July 2001.

Schuster A (1898). On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. Terrestrial Magnetism Atmospheric Electricity 3:13–41.

Schwender D, Madler C, Klasing S, Peter K, Poppel E (1994). Anesthetic control of 40-Hz brain activity and implicit memory. Consciousness Cognition 3:129–147.

Serafini G, Brizi S (1998). Usefulness of auditory steady-state responses (SSR 40Hz) in post trauma coma patients. In: Litscher G, ed. Monitoring of stimulus-induced 40Hz brain oscillations. Scottsdale: Pabst Science Publishers, pp. 215–223.

Spydell JD, Pattee G, Goldie WD (1985). The 40 Hertz auditory event-related potential: normal values and effects of lesions. Electroencephalogr Clin Neurophysiol 62:193–202.

Stapells DR, Linden D, Suffield JB, Hamel G, Picton TW (1984). Human auditory steady state potentials. Ear Hear 5:105–113.

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

Stapells DR, Galambos R, Costello JA, Makeig S (1988). Inconsistency of auditory middle latency and steady-state responses in infants. Electroencephalogr Clin Neurophysiol 71:289–295.

Stapells D, Picton T, Durieux-Smith A (1994). Electrophysiologic measures of frequency-specific auditory function. In: Jacobson JT, ed. Principles and applications of auditory evoked potentials. New York: Allyn and Bacon, pp. 251–283.

Stefanatos GA, Green GG, Ratcliff GG (1989). Neurophysiological evidence of auditory channel anomalies in developmental dysphasia. Arch Neurol 46:871–875.

Stefanatos GA, Foley C, Grover W, Doherty B (1997). Steady-state auditory evoked responses to pulsed frequency modulations in children. Electroencephalogr Clin Neurophysiol 104:31–42.

Steriade M, Amzica F, Contreras D (1996). Synchronization of fast (30–40 Hz) spontaneous cortical rhythms during brain activation. J Neurosci 16:392–417.

Steriade M, Timofeev I, Durmuller N, Grenier F (1998). Dynamic properties of corticothalamic neurons and local cortical interneurons generating fast rhythmic (30–40 Hz) spike bursts. J Neurophysiol 79:483–490.

Stürzebecher E, Cebulla M (1997). Objective detection of auditory evoked potentials. Comparison of several statistical tests in the frequency domain on the basis of near-threshold ABR data. Scand Audiol 26:7–14.

Stürzebecher E, Cebulla M, Wernecke K (1999). Objective response detection in the frequency domain: comparison of several q-sample tests. Audiol Neurootol 4:2–11.

Stürzebecher E, Cebulla M, Pschirrer U (2001). Efficient stimuli for recording of the amplitude modulation following response. Audiology 40:63–68.

Suzuki T, Hirabayashi M (1987). Age-related morphological changes in auditory middle-latency response. Audiology 26:312–320.

Suzuki T, Kobayashi K (1984). An evaluation of 40-Hz event-related potentials in young children. Audiology 23:599–604.

Szyfter W, Dauman R, de Sauvage RC (1984). 40 Hz middle latency responses to low frequency tone pips in normally hearing adults. J Otolaryngol 13:275–280.

Tang Y, Norcia AM (1993). Improved processing of the steady-state evoked potential. Electroencephalogr Clin Neurophysiol 88:323–334.

Tang Y, Norcia AM (1995). An adaptive filter for steady-state evoked responses. Electroencephalogr Clin Neurophysiol 96:268–277.

Tennigkeit F, Ries CR, Schwarz DW, Puil E (1997). Isoflurane attenuates resonant responses of auditory thalamic neurons. J Neurophysiol 78:591–596.

Thornton C, Barrowcliffe MP, Konieczko KM, et al (1989). The auditory evoked response as an indicator of awareness. Br J Anaesth 63:113–115.

Valdes JL, Perez-Abalo MC, Martin V, et al (1997). Comparison of statistical indicators for the automatic detection of 80 Hz auditory steady state responses. Ear Hear 18:420–429.

Van der Reijden CS, Mens LHM, Snik AFM (2001). Comparing signal-to-noise ratios of amplitude modulation following responses from four EEG derivations in awake normally hearing adults. Audiology 40:202–207.

Vander Werff KR, Brown CJ, Gienapp BA, Schmidt Clay KM (2002). Comparison of auditory steady-state response and auditory brainstem response thresholds in children. J Am Acad Audiol 13: 227–235.

Victor JD, Mast J (1991). A new statistic for steady-state evoked potentials. Electroencephalogr Clin Neurophysiol 78:378–388 [published erratum appears in Electroencephalogr Clin Neurophysiol 1992;83:270].

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

Woldorff MG, Hansen JC, Hillyard SA (1987). Evidence for effects of selective attention in the middle latency range of the human auditory event-related brain potential. Electroencephalogr Clin Neurophysiol Suppl 40:146–154.

Wong WYS, Stapells DR (2001). Brainstem and cortical mechanisms underlying the binaural masking level difference in humans: an auditory steady state response study. Paper presented at 17th meeting of the Evoked Response Audiometry Study Group, Vancouver, Canada, July, 2001.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 11: 411–423.

Zar JH (1999). Biostatistical analysis, 4th edn. Upper Saddle River: Prentice Hall.

Zurek PM (1992). Detectability of transient and sinusoidal otoacoustic emissions. Ear Hear 13:307–310.

# M·A·S·T·E·R

### Multiple Auditory Steady-STate Evoked Response Technology

# Your First Choice for ASSR Testing

## Bio-logic Family of Products
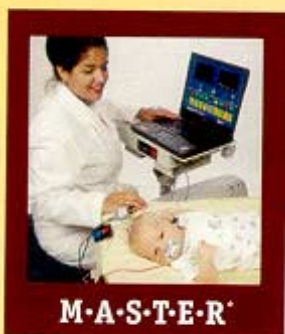
**AEP**
Standard auditory evoked potentials including ABR, ECochG, MLR, ALR, P300 and EABR

**M·A·S·T·E·R***
Multiple Auditory Steady-State Evoked Response Technology

**ABaer®**
Infant hearing screening system – automated ABR, or OAE or both!

**Stacked ABR™**
Advanced technology to assist in the assessment of auditory nervous system abnormalities

**SCOUT®**
Full range otoacoustic emission test capability including DPOAE, TEOAE and SOAE

*Developed in cooperation with Sasha John, Ph.D. and Terence W. Picton, Ph.D. at Rotman Research Institute of Baycrest Centre, Toronto, Ontario, Canada. Patent applied for.
**Developed in cooperation with Manuel Don, Ph.D. at House Ear Institute. Patent Numbers: 6080112 and 6264616B1

To complement our Navigator Pro family of systems, Bio-logic also offers the following portable hearing screening systems:

**AuDX®**
Portable DPOAE or TEOAE testing for all ages

**ABaer® CUB**
PDA-based infant hearing screener, automated ABR or OAE or both

## Please Contact your local dealer

**Bio-logic**
SYSTEMS CORP