Hearing in Time: Evoked Potential Studies of Temporal Processing

Terence Picton

This article reviews the temporal aspects of human hearing as measured using the auditory evoked potentials. Interaural timing cues are essential to the detection and localization of sound sources. The temporal envelope of a sound—how it changes in amplitude over time—is crucially important for speech perception. Time is taken to integrate, identify, and dissolve auditory streams. These temporal aspects of human hearing can be examined using the auditory evoked potentials, which measure the millisecond-by-millisecond activity of populations of neurons as they form an auditory percept. Important measurements are the time taken to localize sounds on the basis of their interaural time differences as measured by the cortical N1 wave, the contribution of the vocal cord frequency and phonemic frequency to the perception of speech sounds as indicated by the envelope-following responses, the temporal integration of sound as assessed using the steady state responses, and the duration of auditory memory as shown in the refractory periods of the slow auditory evoked potentials. Disorders of temporal processing are a characteristic feature of auditory neuropathy, a significant component of the hearing problems that occur in the elderly, and a possible etiological factor in developmental dyslexia and central auditory processing disorders. Auditory evoked potentials may help in the diagnosis and monitoring of these disorders.

(Ear & Hearing 2013;34;385–401)

INTRODUCTION

Time provides a fundamental dimension for perceiving sounds (Viemeister & Plack 1993; Eddins & Green 1995; Phillips 2012). Temporal processing is essential for the detection of sounds, their localization in space, the evaluation of their periodicity, and their identification on the basis of when and how they change. This article considers the temporal aspects of human hearing as measured using the auditory evoked potentials (AEPs). The trigger for this review was a meeting to celebrate the scientific career of Arne Starr. Some of his most highly cited articles (e.g., Starr et al. 1996) deal with the disorder known as auditory neuropathy, which has a devastating effect on temporal processing. This review is not exhaustive because there are too many waves and too little time. In areas with many articles, only those most recent and most representative are considered. This review concentrates on the electroencephalographic (EEG) responses because these are readily recordable, and also because they can record brainstem activity. However, studies using magnetoencephalography (MEG) are also mentioned, because these provide more accurate measures of activity in the auditory cortex.

Human AEPs are primarily categorized in terms of their latency. Early or fast responses come from the cochlea and brainstem (the auditory brainstem response [ABR]); the middle-latency responses derive from the initial activation of auditory cortex; late or slow responses come from auditory and association cortices. AEPs can also be classified by how they respond over time. Transient responses are evoked by stimulus changes, whereas sustained responses last through the duration of a stimulus. Following responses are somewhere between transient and sustained: they are evoked by a stimulus that changes repetitively. Following responses can track the frequency of a sound (frequency-following responses or FFRs) or the envelope of a modulated sound (envelope-following responses). If the changes in the sound are periodic, the following response becomes an auditory steady state response (ASSR).

All these different responses have been evaluated in relation to temporal processing. The upper half of Figure 1 shows the transient AEPs recorded at different latencies. The lower half shows the following responses recorded at different stimulus rates. The following responses are plotted differently from the transient responses, using the root mean square amplitudes over a period of time, rather than the amplitude at a specific time. These amplitudes never go below zero. For the following responses the x axis represents frequency rather than time. Furthermore, values on the axis increase from right to left to allow easier comparison between homologous transient and following responses. The estimated amplitudes for the following responses derive from the findings in multiple articles (Rees et al. 1986; Picton et al. 1987; Maiste & Picton 1989; Cohen et al. 1991; Purcell et al. 2004; Dajani & Picton 2006; Alaerts et al. 2009). The graphs are reasonably well established for the middle and fast frequencies. However, the following responses at slow rates are variable from subject to subject and change with the different stimuli that have been used to evoke them. Most articles report responses with peak amplitudes near 4, 10, and 20 Hz. These have been labeled θ, α, and β in keeping with the frequency bands of the human EEG.

FREQUENCY AND PERIODICITY

The perceived frequency of a sound derives from two separate physiological mechanisms. The first is von Békésy's travelling wave, which causes a particular region of the basilar member to respond best to a particular frequency. The frequency of a sound is thereby coded in terms of the place of maximal activation on the basilar membrane, with high frequencies coming from the basal turn and low frequencies from the apical turn. The central nervous system relates the source of its activation to a location on the basilar membrane by means of topographic maps and labeled lines. The second mechanism for perceiving frequency is the ability of afferent neurons (or populations of neurons) to lock themselves to a particular phase of the sound. The frequency of the sound is then coded in terms of time. The central nervous system estimates either the interdischarge interval or the autocorrelation function of incoming neuronal activity. Timing codes are only possible for sounds with frequencies less than 2000 Hz. At higher frequencies reliable neuronal phase locking does not occur.

University of Toronto, Toronto, Ontario, Canada.

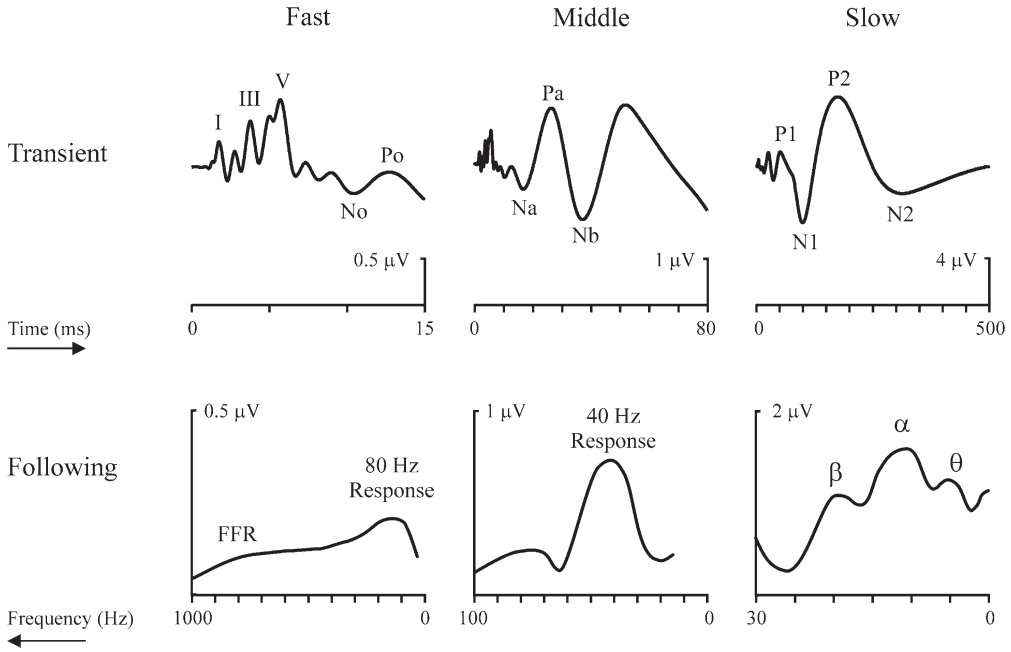

Fig. 1. Transient and following responses. The upper half shows the transient auditory evoked potentials plotted on three different time scales. The waveforms represent the typical response to 70 dB nHL clicks presented at a rate of 1 per sec. (see Picton, 2011, Fig. 2-2). The lower half represents the following responses that are recorded to amplitude-modulated noise when the rate of modulation is varied. The plotted amplitudes represent the typical response to amplitude modulated white noise presented at approximately 60 dB SPL. These following responses are plotted on three different frequency scales with the slower responses on the right. Though initially counterintuitive, this arrangement of the axis allows comparison between homologous transient and following responses. The peaks of the slow following responses near 4, 10, and 20 Hz have been named θ, α, and β after the frequency bands of the human EEG. EEG, electroencephalographic.

Frequency-Following Responses

The hair cells on the basilar membrane can follow the frequencies of sounds up to the limits of human hearing. In doing so, they generate the cochlear microphonic (CM). Afferent neurons can follow the sounds exactly only up to their maximum discharge rate of several hundred hertz. However, single neurons may be able to respond on every second, third, or fourth cycle of the sound. Populations of neurons may thus be able to follow sounds up to frequencies of about 2000 Hz and generate a "neurophonic." This neurophonic can derive from both primary afferent neurons (Snyder & Schreiner 1984) and neurons in the auditory brainstem (Smith et al. 1975). Both the microphonic and the neurophonic can be recorded from electrodes on the human scalp or mastoid. The largest of the human scalp–recorded responses, occurring with a delay of about 6 msec on vertex to mastoid recordings, is called the "frequency-following response" or FFR (reviewed by Krishnan 2007). This scalp-recorded response is generated in several parts of the auditory pathway: pontine regions (best recorded using a mastoid to mastoid montage) respond with a latency of 2 to 3 msec; and midbrain regions (best recorded using a vertex to neck montage) with a latency of 5 to 7 msec. Early studies of this response evaluated its possible use in the assessment of low-frequency hearing thresholds, but this is now performed with tone-ABRs or ASSRs.

The recent resurgence of interest in the FFR is related to the possibility that it might evaluate temporal processing in the brainstem. Kraus and her colleagues (Johnson et al 2005; Skoe & Kraus 2010) have proposed the use of a brief speech stimulus (da) that can evoke both an onset ABR (to the beginning of the stimulus) and an FFR (to the vowel frequencies). The FFR has been reported to be enhanced by attention (Galbraith et al. 1998). It is also increased by experience and training, being larger in subjects who speak tonal languages compared with those who do not (Krishnan et al. 2004, 2005), and in musicians compared with nonmusicians (Wong et al. 2007, Bidelman et al. 2011). Clearly, the FFR represents the temporal information present in a sound and as such, might be involved in the analysis of periodicity and pitch. However, the exact relationship of the FFR to pitch perception remains to be determined because the FFR may not specifically represent all the perceived pitch information in complex stimuli (Gockel et al. 2011). Although the temporal parameters of sounds need to be represented and analyzed, the final perception of pitch requires a cognitive decision about how these temporal parameters can be best interpreted.

Rippled Noise

Rippled noise is a stimulus that has been added to itself after a delay, like the combination of a source with its echo. The resultant sound shows some temporal regularity, which can be demonstrated by peaks in its autocorrelation function. This is perceived as a pitch equal to the reciprocal of the delay. The power spectrum shows enhancements ("ripples") at harmonics of this basic frequency. The perceived pitch becomes more prominent when the number of iterations used in making the rippled noise is increased. Psychophysical studies have indicated that this perceived pitch is determined by the temporal regularity rather than the spectrum of the sound (Yost 1996).

Iterated rippled noise can evoke human brainstem FFRs (Swaminathan et al. 2008) and these have been used to study

the effects of linguistic and musical experience (e.g., Bidelman et al. 2011). One advantage of using rippled noise instead of tones or formants is that the effects of signal to noise levels can be studied by varying the number of iterations used to generate the stimulus.

MEG recordings show that the onset of a rippled noise elicits both a cortical N1 response and a later negative wave (Krumbholz et al. 2003). The initial N1 is likely a response to the onset of the sound, whereas the later negative wave may register the pitch of the stimulus. If the stimulus is a change from a simple broadband noise to a rippled noise only the later wave is present. Krumbholz et al. (2003) reported the peak latency of their "pitch-onset response" as equal to 120 + 4d msec, where d is the delay parameter of the rippled noise. The increased latency of the response also occurs in the EEG responses to the simple onset of rippled noise (as compared with the onset of a complex harmonic sound) (Butler & Trainor 2012). In these electrical responses, the N1 to the onset of a sound probably overlaps the later negative wave related to the extraction of the pitch information.

Changes in the parameters of an ongoing rippled noise also evoke just the late cortical response (Hertrich et al. 2004). More recent studies have shown similar responses to changes in rippled noise using EEG recordings (Won et al. 2011). The negative wave evoked by the change in the noise peaks later than that elicited by the onset of a sound or by a change in the frequency of a tone. Hertrich et al. (2004) reported a peak latency of 136 msec (using rippled noise with a delay of 9 or 7.5 msec) compared with 110 msec for the onset of a click train. The longer latency of the pitch response compared with a simple onset response likely represents the increased time required for processing the pitch of the sound and determining that it differs from the pitch of the preceding sound.

SOUND LOCALIZATION

The most important way that we localize a sound source is by comparing the timing and intensity of the sounds reaching our two ears. Monaural localization processes based on the acoustics of the pinna are minimal compared with such binaural processing. Interaural time differences (ITDs) are of two kinds. For transient stimuli—brief sounds or the onsets of longer sounds—we can assess which ear receives the sound first and by how much. For continuous sounds, we can measure the ongoing ITD (equivalent in the case of pure tones to the interaural phase difference). ITDs for continuous sounds can only be perceived if the auditory system can accurately follow the frequencies within the sound. Generally this means frequencies of less than 1500 Hz, although the cutoff varies with several parameters (e.g., age). Our ability to distinguish ITDs is extremely precise. We can discriminate the locations of two transient sounds located in front of us, even when their ITDs differ by less than 90 µsec (Yost et al. 1971; Dingle et al. 2010). A sound directly ahead has an ITD of 0 µsec and one a few degrees to one side or the other will have an ITD of several tens of microseconds. For longer-lasting tones, ITDs of 10 µsec can be discriminated (Mills 1958; Yost 1974).

Localization of Transients

Many studies have looked at the AEPs to binaural clicks. A simple approach is to determine the difference between the responses to a binaural click and to a monaural click. The binaural stimulus is perceived as one stimulus (in the center of the head) rather than two stimuli (one in each ear). This "binaural fusion" is registered in the brainstem. However, there is only a small difference between the ABR to a binaural stimulus and the sum of the monaural responses recorded when the stimuli are presented to each ear separately. The small difference between the binaural waveform and the summed monaural responses is studied as the "binaural interaction component" (Dobie & Norton 1980).

The effects of binaural fusion are greater in the cortical responses (both middle-latency and slow components) where the response to a binaural stimulus is similar to the response to a single monaural stimulus (McPherson & Starr 1993). The cortex apparently processes sounds according to their locations in space rather than the ear or ears in which the sounds are received (Picton & Ross 2010).

More complex physiological evaluations of localization of transient assess the effects of changing the ITD. McPherson and Starr (1995) found that the binaural interaction component increased in latency and decreased in amplitude as the ITD increased. These and similar data have been interpreted by Riedel and Kollmeier (2006) in terms of different models of how the brainstem auditory neurons track interaural timing. The ABR data can then be used to estimate the timing of excitatory and inhibitory inputs to binaural neurons. The model of Ungan et al. (1997), for example, fits the recorded measurements of the binaural interaction component if the inhibitory input to the binaural neurons arrives slightly earlier than the excitatory input.

In a reverberant environment, the sounds from an auditory source reach the ear both directly and indirectly as echoes reflected off nearby surfaces. The perceived location of the sound is dominated by the interaural characteristics of the initial input—the "precedence effect." Liebenthal and Pratt (1999) recorded the AEPs to binaural clicks that were followed by an echo (from a different location). By subtracting the AEP to the click alone from the response to the click and echo, the authors could assess the response to the echo. The ABR to the echo was essentially normal, indicating that the auditory brainstem processed both the initial sound and its echo. However, the Pa wave of the middle-latency response in response to the echo was reduced, suggesting that echo suppression occurs at the level of the auditory cortex where Pa is generated. Damaschke et al. (2005) confirmed the lack of any precedence effect in the brainstem and postulated that the perceived location of a sound in a reverberant environment depends on cortical processing.

Localization of Continuous Sounds

Sound localization can also be studied using continuous rather than transient sounds. The simplest stimulus is a change in ITD. For a broadband noise, this change cannot be heard monaurally and any response indicates binaural temporal processing. Changes in interaural intensity could be heard as decrements or increments in monaural intensity.

Two steps are involved in the localization of continuous sounds on the basis of ITD. First, the auditory system must realize that the sounds arriving at each ear are sufficiently correlated so that a consistent ITD can be measured. Changing dichotic noise from uncorrelated (completely different stimulus in each ear) to correlated (same stimulus in each ear) is heard as a diffuse noise suddenly becoming focused at the midline. This

stimulus evokes a cortical N1 wave with a peak latency that is approximately 30 msec later than the N1 to a simple onset (Jones et al. 1991; Dajani & Picton 2006).

Second, the ITD must be measured to place the stimulus along a dimension that goes between left and right. If the stimulus is the same in each ear but changes its ITD from left-ear-leading to right-ear-leading, the sound is perceived as moving suddenly from left to right. This stimulus evokes an N1 wave with a peak latency approximately 35 msec later than that to stimulus onset (Halliday & Callaway 1978; McEvoy et al. 1990, 1991). These results are illustrated in Figure 2. This response can be recognized with ITD changes as small as ±100 μsec, although, accurate physiological thresholds have not been assessed (McEvoy et al. 1991).

The ITD transition is very fast. For a normal adult head, the time lag from left ear to right ear varies between ±0.7 msec (Moore 2003, Chapter 7), although, we can still fuse and lateralize sounds with lags of several milliseconds. Despite its sudden onset, however, the time taken to identify the ITD change is longer than for a simple stimulus onset. In psychophysics, this extra time is described in terms of "bilateral sluggishness" (Blauert 1972; Grantham & Wrightman 1978; Grantham 1995). Recent physiology has indicated that this slowness is not evident in the brainstem auditory neurons, which rapidly and accurately track changes in the interaural correlation and latency delay (Joris et al. 2006; Siveke et al. 2008). It must therefore derive from the cortical assessment of the information arriving from the brainstem. Binaural sluggishness might therefore be demonstrated in the slow AEP as the delayed latency of the N1 for ITD changes compared with the N1 to stimulus onsets.

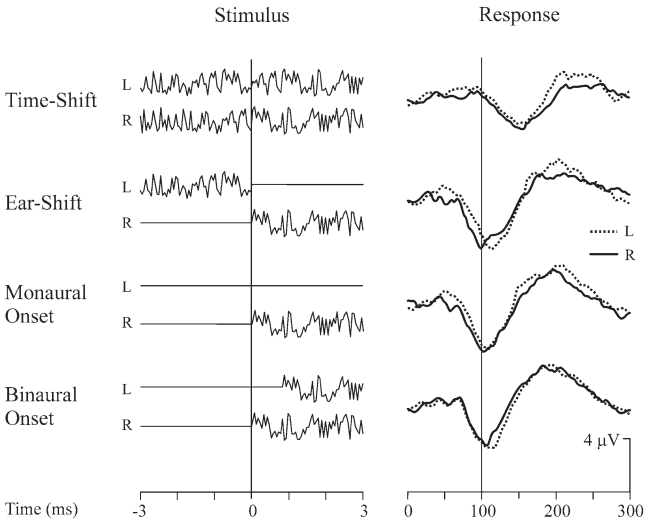

Fig. 2. Evoked potentials to a shift in lateralization of a sound. The first set of data show the responses to a change in the interaural time differences of a dichotic noise from the left-ear-leading by 0.7 msec to the right ear leading by 0.7 msec (or vice versa). The second set of data show the responses to a simple change in the ear of stimulation. The third shows the responses to the onset of a stimulus in one ear, and the fourth shows the response to the onset of a binaural stimulus with one ear leading the other. The stimuli illustrated are for the stimuli going from left to right or just beginning on the right. The responses are shown for both these stimuli (R) and for those shifting to or beginning on the left (L). All responses show a large N1-P2 waveform. The responses to the time-shift stimulus are significantly later than the others. Data from McEvoy et al. (1990).

Another interesting aspect of the evoked potentials to these binaural stimuli, either the shift from one ITD to another, or the shift from one level of interaural correlation to another, is the absence of any clear early or middle responses. Nothing happens until the P1 wave of the slow AEP. Prolonged recordings to reduce the residual EEG noise failed to show any clear early transient AEPs, or at least, none above the noise levels of approximately 0.1 to 0.2 μV (McEvoy et al. 1990)—why exactly is a matter for speculation. The brainstem auditory system may calculate a running measurement of the interaural correlations at different ITDs and simply provide this ongoing information to the cortex, where a decision is then made that a stimulus has changed. Another possibility is that the decision occurs in the brainstem but with a variable latency because of the continually cycling nature of the correlation process. If the correlation and ITD measurements are based on a window of 25 msec or more, the trial-to-trial latency-jitter in the timing of the decision would cause the average brainstem and middle-latency evoked potentials to cancel themselves out. The slow onset response in the cortex would not be cancelled but would be delayed and attenuated by this amount of jitter (because the positive waves partially overlap with negative waves and vice versa during averaging).

A sudden change in the interaural timing (or phase) of a pure tone rather than a noise, causes an audible transient. The nature of a response to such a change is unclear, because it might be evoked by the audible click rather than the change in phase. Ross et al. (2007b) proposed a simple method to eliminate the transient: the stimulus was amplitude modulated so that the change in phase occurred when the amplitude of the stimulus was zero. The change from in to out-of phase is heard as an increase in the spaciousness of the tone, a pseudostereo effect. The change evokes a P1-N1-P2 response with a latency that is about 25 msec later than for the response to sound onset. In young subjects, a clear response occurs for carrier frequencies up to 1500 Hz.

This stimulus also has the advantage that it also allows us to monitor the ASSR evoked by the modulation of the tone. When the phase changes, the ASSR rapidly decreases in amplitude and then slowly becomes reinstated. This transient decrease in the response can be used to estimate the threshold for recognizing a phase change with the same sensitivity as the N1-P2 response. Its advantage is that it can be recorded at faster rates. The N1-P2 gets smaller as the interval between stimuli decrease below several seconds but the change in the steady state response can be recorded at rates of several stimuli per second (Ross 2008; Picton & Ross 2010).

The negative wave in the responses to changes in coherence, ITD, or interaural phase is likely the same as the N1 response to the onset of a sound but delayed by the time required to assess the interaural correlation between the stimuli. Another explanation is that it represents a mismatch negativity (MMN) evoked by a change from a previous stimulus (e.g., Jones et al. 1991). However, the N1 and MMN may indicate a similar cortical response to change (from nothing or from a prior stimulus), with each response involving areas of cortex specific to the type of change (discussed at greater length in Picton 2011, chapter 11).

Binaural Beats

When two sinusoidal signals of slightly different frequencies are combined they form beats with a frequency equal to

the difference between their frequencies. If the signals are sent to separate ears, "binaural beats" can be heard in the midline even though there are no actual acoustic beats. The perceptual phenomenon only occurs for stimuli with frequencies less than 1500 Hz, suggesting brainstem interactions between the temporal representations of the stimuli. Physiological responses to binaural beats can be effectively recorded as an ASSR if the beat frequency is set near 40 Hz (Schwarz & Taylor 2005). If the beat frequency is set at 3 or 6 Hz, cortical responses can be recorded to each beat (Pratt et al. 2010). Because the beat rate is faster than the normal rate for eliciting the slow cortical AEPs, the responses are small and consist largely of the P1 wave. This wave was smaller and peaked later (by an average of 44 msec) for the response to binaural beats compared with the response to acoustic beats.

Larger cortical responses to binaural beats can be obtained using an elegant new approach (Ozdamar et al. 2011). Continuous tones of the same frequency and opposite polarity are presented binaurally. At a rate of 1 per second, the tone in one ear is increased for a duration of 20 msec by 20 Hz (without any phase jumps in the signal) and the tone in the other ear decreased by the same amount. This results in a 20 msec (10 msec rise and 10 msec fall) pulsatile binaural beat. The slow AEPs evoked by this stimulus contain responses to the monaural frequency changes and the response to the beat. An approach similar to that used to measure the binaural interaction component can then be used to distinguish the beat response. Separate responses are recorded to the beat and to each of the monaural frequency changes. The response to the beat less the sum of the responses to the monaural frequency changes gives the true binaural beat response. This shows large P1-N1-P2 components with the peak latency of the N1 occurring about 15 msec later than the N1 to the monaural frequency change.

TEMPORAL RESOLUTION

Temporal resolution is the ability to detect changes in the amplitude or spectral content of a sound over time (Viemeister & Plack 1993). Changes in the amplitude of a sound without concomitant change in its spectral content or localization form the envelope of the stimulus "within" a sensory channel. If the sounds change frequency or localization, then we must monitor changes "across" sensory channels. The time taken to detect changes across channels is typically between 5 and 10 times greater than the time taken within a channel. However, there is little correlation between the performances on the two tasks, suggesting that they derive from separate neural mechanisms (Phillips & Smith 2004). Across-channel monitoring is much more variable from subject to subject and likely requires perceptual training and significant cognitive effort. Across-channel gap detection thresholds likely share the same perceptual mechanisms as the categorical perception of voice-onset times (Elangovan & Stuart 2008).

Three common ways to assess our ability to follow the envelope of sounds are the detection of transient gaps in sounds, the recognition of double versus single stimuli, and the discrimination of amplitude-modulated from unmodulated sounds. The three approaches are related: a gap is essentially a single cycle of modulation, and a gap occurs between the components of a double stimulus. Gap detection, itself, can be examined in two ways: either the detection of an occasional gap in an ongoing sound or the discrimination of a sound containing a gap from one without any gap.

Following the temporal fluctuations in the amplitude-envelopes of ongoing sounds is essential to their comprehension. The envelope of a sound is as important as its spectral content in the determination of the pitch and timbre of musical sounds (e.g., Rasch & Plomp 1999). Most of the discriminations necessary for speech processing involve changes across channels, for example, between the onset of the sound and onset of voicing, or between the frequencies at the beginning of a stop consonant and the onset of the vowel. The envelopes carry much of the information that we use to understand speech sounds, especially when separate envelopes are determined for different frequency bands (Shannon et al. 1995).

Gap Detection

The onset of the sound at the end of a gap evokes a clear ABR with a latency similar to the normal onset response (Poth et al. 2001; Werner et al. 2001). Using gaps in broadband noise Werner et al. (2001) found that the average gap detection threshold (2.4 msec) as assessed using the ABR was similar to the average psychophysical threshold (2.9 msec) although the correlation between the two thresholds was only 0.39.

Cortical N1 and P2 responses are also evoked by an occasional gap in a continuous tone (Harris et al. 2012; Michalewski et al. 2005) or by the gap separating two brief sounds (Ross et al. 2010). Harris et al (2012) did not specifically assess gap detection thresholds but found that the detection of brief gaps in continuous noise were similar using either cortical AEPs or behavioral responses.

When gaps occur in an ongoing sound, AEPs should theoretically be elicited by both the offset and the onset of the sound. However, offset responses are smaller than onset responses and a separate offset response only becomes apparent when the gap is 50 msec or more. If the timing is referred to the beginning of a gap lasting 50 msec, two negative waves can be distinguished—at latencies 90 and 150 msec—representing the N1 response to the offset of the sound, and the N1 response to the onset of the sound after the gap (Michalewski et al. 2005, their Figure 4).

The detection of a gap in an ongoing sound occurs independently of sound localization. A gap in a sound that lags behind another sound is still heard (and evokes an N1-P2 response) even though the lagging sound is not (Li et al. 2005). The gap is perceptually "captured" by the leading sound.

Double Stimuli

Determining whether one or two stimuli have occurred is similar in many ways to gap detection. In the single–double discrimination, the subject responds to one or two stimuli rather than to a stimulus containing a gap or one without.

The late AEPs to pairs of stimuli have been most commonly studied at intervals when it is perceptually obvious that there is a second stimulus. The focus is on how the brain responds to the second stimulus rather than on thresholds for detecting a double stimulus. The late AEPs to the second of a pair of stimuli decrease in amplitude as the interval between the stimuli decreases from 10 to 0.5 sec. However, at intervals shorter than 0.5 sec there is an enhancement of the response, so that between 100 and 400 msec the second stimulus evokes a response that is equal to or even larger than the first. This was initially

recognized in MEG studies (e.g., Loveless et al. 1996) where the response components following the N1 to the first stimulus are small and do not significantly overlap with the response to the second stimulus. The effect can also be seen in electrical recordings, provided the response to the initial stimulus is subtracted from the responses to the paired stimuli (Budd & Michie 1994; Sable et al. 2004).

Several mechanisms have been proposed to explain this enhancement. Loveless and his colleagues (1996; McEvoy et al. 1997) found that their results could be best explained using two separate sources for the N1 with the anterior source (occurring approximately 30 msec later than the posterior source) contributing most to the enhancement effect. This anterior N1 source is related to identifying the nature of a stimulus (answering the question "what?") whereas more posterior source is related to localizing the stimulus ("where?") (Jääskeläinen et al. 2004). Sable et al. (2004) propose that the prolonged refractoriness of the N1 response is mediated by inhibitory processes that take several hundred milliseconds before they develop. Responses to stimuli occurring during the period when the inhibition remains latent are therefore not reduced in amplitude like later responses. Another possibility might be that an MMN is evoked by the second stimulus, which can be perceived as deviant if the memory representation of the preceding stimulus has not had enough time to form (Wang et al. 2008).

Comparing the responses with single and double stimuli is difficult in children because their late AEP has a different morphology from that of adults (Picton & Taylor 2007; Picton 2011, chapter 11). The major component of their late AEP is a large P1 wave, with the N1 being very small and only detectable at long interstimulus intervals (ISIs). The normal adult N1 response does not develop until early adolescence. Fox et al. (2010, 2012) have developed a method to determine the presence of the response to the second of two stimuli that is independent of the morphology of the response. The response to the double stimulus is compared with the response to the single stimulus by calculating the correlation between the responses and comparing this with the distribution of correlations expected when there was no second stimulus. A significantly low correlation means that the responses are different, because there was an extra response to the second stimulus. In this way they were able to show that normal adults have a recognizable response to the second of two tones with intervals as short as 25 msec. In children of 7 to 9 years of age a response to the second tone was only distinguishable at intervals of 200 msec or more. The threshold for showing a response to the second tone correlated with the child's performance on nonword repetition and decreased as the child became older. The developing ability to detect the second stimulus might therefore provide the basis for improving speech processing.

The late AEPs to consonant–vowel syllables often suggest overlapping responses to the consonant and the vowel (Tremblay et al. 2003). For some syllables such as "shi," the double response is particularly prominent, as shown in Figure 3, where the peaks in the response to the vowel onset are indicated by the asterisks (data from Friesen & Picton 2010). There may be some relationship between the particular sensitivity of the N1-P2 system to intervals between 50 and 400 msec and the temporal structure of normal speech. Such intervals are similar to the durations of phonemes and syllables.

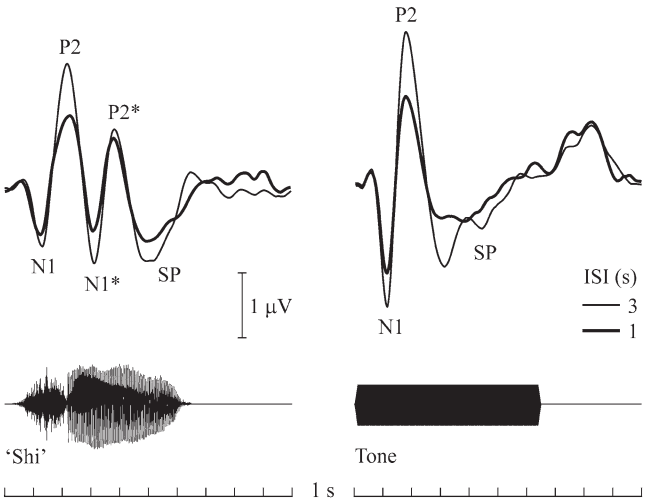

Fig. 3. Evoked potentials to phonemic changes. The responses in the left half of the figure were evoked by the syllable "shi." They demonstrate an N1-P2 response to the onset of the stimulus, a second response (asterisks) to the onset of the vowel, and a SP to the continuation of the vowel. On the right are shown the auditory evoked potentials to a 1 kHz tone, which contains only the N1-P2 response to the onset of the tone and the SP to its continuation. Responses recorded with ISIs of 3 sec are significantly greater than the responses at ISIs of 1 sec. Data from Friesen and Picton (2010). SP, sustained potential; ISI, interstimulus interval.

Envelope-Following Responses

The ability of the human brain to respond to rapidly changing sounds can be studied in two ways. The simplest is just to record the steady state responses to stimuli presented at different rates (e.g., Rees et al. 1986; Picton et al. 1987). The second is to track the response to a sweeping stimulus rate (Purcell et al. 2004). Both techniques show a clear enhancement of the response at frequencies near 40 Hz (Picton 2011). The lower half of Figure 1 shows the amplitude of following responses at different stimulus rates.

At higher frequencies, responses are again particularly evident at 80 to 120 Hz and then fall off toward higher frequencies with little response evident above 1500 Hz. Whatever responses do occur at these high frequencies, they likely originate in the cochlea (microphonics) rather than the brain (neurophonics). The brain responses to envelopes of 80 to 300 Hz might be important in synchronizing the spectral analysis of voiced sounds. Voiced speech is far more comprehensible and resistant to noise than whispered (unvoiced) speech.

Following responses at slow rates have been more widely studied in recent years. These responses vary with the stimulus used (amplitude modulation, frequency modulation, etc.) and the pattern shown in the lower right of Figure 1 is only rough approximation of what might be recorded. A large response tends to occur in the α frequency range of the electroencephalogram (8–13 Hz). Often, this response occurs at twice the frequency (second harmonic) of modulation. For example, Maiste and Picton (1989) recorded a large response at 8 Hz to a sound that was frequency modulated at 4 Hz. Dajani and Picton (2006) found a similar 8 Hz response to 4 Hz cycles between correlated and uncorrelated dichotic noise. Such second-harmonic responses must be generated by a system that reacts to a change in the stimulus (in frequency or in the amount of correlation) regardless of the direction of the change.

AUDITORY SCENE ANALYSIS

Our perception of the auditory world has been described by Bregman (1990) as a process of "auditory scene analysis" (recent reviews by Shamma & Micheyl 2010; Snyder et al. 2012). We perceive the visual world on the basis of various rules, such as things that move together tend to be part of a single object. The auditory system follows its own rules. The categorization of auditory sounds into "objects" is largely based on localization: sounds coming from one point in space tend to be part of a single auditory object. However, spectral patterns also play a role: a complex sound coming from a single natural source is typically composed of a set of harmonics of a single fundamental frequency. Harmonics perceived as tuned to a single fundamental tend to be perceived as a single auditory object.

If we attend to sounds over a period of time, we can separate different sounds coming from a similar location into "streams," grouping some sounds together and separating others on the basis of spectral relations. The temporal characteristics that determine streaming are varied: different sounds modulated by the same envelope can be streamed together, as can different sounds that follow temporal or rhythmic patterns. The process of auditory stream analysis involves both bottom–up and top–down processes (Alain et al. 2001b). The physical characteristics of the sounds determine how they might be perceived as objects or streams. However, what is finally perceived involves interactions between sensory information and attentional predictions (Alain & Winkler 2012).

Auditory Objects

A simple paradigm to assess the recognition of separate auditory objects uses the mistuned harmonic. A sound made up of multiple harmonics of a single fundamental will be perceived at the pitch of the fundamental. If the harmonics are all at the same intensity, the stimulus sounds like a buzz. If the harmonics follow particular patterns determined by the resonance characteristics of a source, they can sound like musical notes. If one of the harmonics is mistuned such that it is no longer a multiple of the fundamental, it will become perceptible as an independent sound object: a pure tone in the buzz, or an aberrant note in the musical sound. The ability of the human auditory system to detect a mistuned harmonic falls off dramatically when the frequency of the harmonic exceeds 2000 Hz. This suggests that the process of detecting the mistuning depends on the ability of the auditory system to synchronize with incoming sounds (Hartmann et al. 1990). The detection process may line up the different frequency components over time so as to determine which continue to be synchronous as the period of evaluation increases.

The slow AEP to a stimulus with a mistuned harmonic differs from that evoked by a stimulus with all harmonics in tune by having a smaller P2 wave. The calculation of a mistuned-tuned difference waveform isolates a small negative wave associated with the perception of the mistuned harmonic: the "object-related negativity" (Alain et al. 2001b).

A similar object-related negativity occurs during the processing of superimposed speech sounds that become more easily discriminated when the differences in the fundamental frequencies of the speakers increase (Alain et al. 2005). The object-related negativity occurs automatically whether or not the subject is actively attending to the stimuli. If the subject is paying attention, the object-related negativity is followed by a P400 wave.

Streaming

The classical paradigm for evaluating streaming involves the perception of a gallop. The paradigm presents sounds of two frequencies in a pattern ABA-ABA-ABA… If the frequencies are close, we perceive a gallop. However, as the frequencies become further apart we perceive two separate streams, each occurring at their own rate: A-A-A-A-A-A and -B- - -B- - -B-. The slow AEPs N1 and P2 become larger as the streams become more distinguishable (Gutschalk et al. 2005; Snyder et al. 2006).

In the gallop paradigm, streaming varies with the timing of the stimuli, the frequency differences between the stimuli, and the attention paid to the stimuli. All three factors affect the N1 wave of the response. The N1 becomes very small when the interval between stimuli decreases, when the frequency difference between the stimuli becomes less than a few semitones, and when the subject does not attend to the individual stimuli. The smaller the amplitude of the N1, the greater the chance that the stimuli are perceived as grouped together. The N1 may thus indicate the occurrence of a stimulus that should be considered separate from that of an ongoing stream.

SPEECH RHYTHMS

The sounds of speech follow many rhythms. The fastest is the rhythm of phonation that occurs during the production of vowels and voiced consonants. In normal speech, an adult male will have a fundamental pitch of approximately 120 Hz, a normal adult female of approximately 200 Hz, and a normal child of approximately 300 Hz. Human song can vary the pitch from approximately 70 Hz to approximately 1500 Hz.

Speech information is carried by a variety of slower rhythms. Phonemes last between 25 to 50 msec for stop consonants and 200 to 300 msec for long vowels, giving rhythms between 3 to 40 Hz. Syllables last between 100 and 500 msec and produce rhythms between 2 to 10 Hz. Words and phrases make up the slower rhythms of prosody or poetry. Our listening brains must tune to all these different rhythms.

Phonation

The basic rhythm underlying much of normal speech is the rate at which the vocal cords vibrate. Responses to the normal voicing rhythm are most easily studied using synthesized speech, where the rate of phonation is held constant and the brain responses analyzed using frequency transforms. However, normal speech does not have a constant voicing frequency. In European languages, the speech fundamental varies both randomly and in response to various linguistic intonation patterns such as the rising pitch at the end of a question. In tonal languages, vowels are categorized on the basis of how the pitch changes during the vowel. Measuring the brain's response to a changing frequency cannot be performed using frequency transforms, which assume that a signal is constant for the duration of the analysis. However, they can be measured using techniques such as short-term autocorrelation (Krishnan et al. 2004) and Fourier analysis (Aiken & Picton 2006, 2008a). High-pass

masking studies show that the brain's response at the voicing frequency is a combination of two processes: an FFR that follows the spectral energy in the signal at the voicing frequency (the spectral FFR) and an envelope-following response that follows the modulation frequency of the higher-frequency harmonics (the envelope FFR) (Aiken & Picton, 2008a). The scalp-recorded envelope FFR is mainly generated in the brainstem. However, MEG recordings have also found sources in the auditory cortices that followed the pitch of spoken words (Hertrich et al. 2012), and intracranial recordings have shown cortical phase-locked field potentials at frequencies up to 200 Hz (Brugge et al. 2009).

Speech Envelopes

The speech envelope is the changing amplitude of the energy in continuous speech as it sequences through the different phonemes and syllables of spoken words. Following this envelope is essential to speech perception. A set of four frequency-limited bands of noise, modulated by the temporal envelopes for four matching frequency bands of the speech signal, can provide enough information to support near-perfect identification of vowels, consonants, and sentences (Shannon et al. 1995). This finding explains why speech can be recognized through a cochlear implant, where the incoming spectrum is divided up across a small number of electrodes rather than across many thousand hair cells. Envelope frequencies between 4 and 16 Hz contribute most to the intelligibility of speech. The faster of these rhythms are related to changing phonemes and the slower to syllable transitions.

Aiken and Picton (2008b) used correlation techniques to look at the response of the human auditory cortex to the speech envelope of sentences. The envelope of a spoken sentence was abstracted from the speech signal by rectification and low-pass filtering. This was then correlated with the averaged EEG response from the auditory cortices using various delays. The highest correlation occurred at a latency of about 190 msec, with the auditory cortex following the speech envelope after this delay. Using MEG recordings and a canonical correlation analysis, Koskinen et al. (2012) found that the delay between speech envelope and cortical activity was approximately 150 msec. These latencies suggests that the slow AEPs (N1 and P2) are likely involved in following speech. The morphology of the transient response might then represent the impulse function of the cortex, which is convolved with the incoming sensory signal to give the ongoing cortical following response. Similar waveforms were obtained by Hertrich et al. (2012) using MEG, and by Lalor and Foxe (2010) using a different mathematical approach ("auditory evoked spread spectrum analysis"). Abrams et al. (2008) found that the cortical response that followed the speech envelope was greater in the right hemisphere than the left.

Compressed speech provides a means for determining whether the auditory system can follow speech at rapid rates. Sophisticated programs can speed up a recorded speech stimulus without altering its main spectral content. The auditory cortex evaluated using either MEG (Ahissar, et al. 2001) or direct intracortical recordings (Nourski et al. 2009) can follow speech sounds at rates up to five times faster than normal (even though comprehension decreases at rates of >3 times and speech becomes essentially incomprehensible at rates of 5 times).

MEG sources in the posterior auditory regions of the temporal lobe follow the amplitude of the fundamental pitch of ongoing natural speech at frequencies near 0.5 Hz (Bourguignon et al. 2013). These slow rhythms may be associated with the parsing of speech into phrases.

We have begun to see how the speech envelope is converted to electrical rhythms in the brain. However, the speech envelope is not a single waveform. Rather, there are multiple envelopes, each for a particular spectral region. To study these envelopes we would have to have some idea of the way in which the auditory system separates frequencies for speech processing.

Furthermore, the processing of the envelope must be more complex than simply representing it in electrical form. Transitions in the envelope (onsets and offsets) are probably the most important parts of the ongoing waveform. Interactions must occur between these envelope transitions and the regions of the brain that process phonemes, syllables, words, and sentences. Hertrich et al (2012) found that triggering the MEG on the onsets was as effective as correlating the MEG with the complete envelope. Pasley et al. (2012) were able to reconstruct the acoustic speech signal from neuronal activity recorded from the posterior temporal lobe, using an algorithm that largely depended on spectral transitions.

As well as locking to the speech signal, the rhythms of activity in the auditory cortex may themselves serve to synchronize processing between different areas. When sentences are perceived, a burst of θ activity occurs at the onset and then continues through the sentence (Howard & Poeppel 2012). How much of this activity is exogenously driven by the speech signal and how much is related to communication between speech processing areas remains to be seen.

TEMPORAL INTEGRATION

The brain must integrate what it has analyzed over a period of time to determine the periodicity of a stimulus, to localize its source using ITD, to identify objects, to separate stimuli into auditory streams, and to follow speech. The human brain uses a variety of different durations (or windows) for these integrations (Eddins & Green 1995). For example, Hirsh (1959) suggested that the auditory system used a window of about 2 msec to recognize a single stimulus, and a longer period of about 20 msec determine when it occurred relative to another stimulus. Longer periods of over 100 msec are required to segregate concurrent events (Divenyi 2004).

Probably the most widely studied paradigm to assess temporal integration considers how stimuli increase in loudness as their duration increases up to approximately 200 msec. At greater durations the stimuli are heard as lasting longer, but not as being louder. However, each of the different types of auditory information may have its own integration time. We have already considered the finding that a period lasting several tens of milliseconds is necessary to detect the periodicity of rippled noise or interaural timing cues.

The ASSR evoked by the 40-Hz amplitude modulation of a tone develops over a period of about 200 ms before continuing as a stable response (Ross et al. 2002). This initial development can be obscured by an overlapping gamma-band response (GBR) to the onset of the modulation: a brief burst of activity with frequencies near 40 Hz. The ASSR and GBR can be distinguished by using stimuli with modulations that begin with opposite

phase, since the GBR follows the onset of modulation regardless of its phase. Calculating the difference between the responses removes the GBR (and other transient responses) but leaves the ASSR. Once developed, the ASSR returns to baseline over about 50 ms when the modulation ceases. If an ongoing 40-Hz ASSR is disrupted, for example, by a concomitant short burst of noise, the ASSR falls off and then takes about 200 ms for the response to develop again (Ross et al. 2005). The 200 ms period that the 40-Hz takes either to develop or reinstate itself after it has been disrupted may reflect some basic temporal integration in cortical processing. The brainstem integrates activity over shorter times. Figure 4 shows some electrical recordings that replicate the 200 ms development time for the 40-Hz response and demonstrate a shorter integration period of about 100 ms for the 80-Hz brainstem response (Picton, 2011, chapter 10).

AUDITORY MEMORY

Incoming sensory information is preserved briefly in a sensory register, typically for several seconds, even when the information is not attended to. In the auditory system, this sensory memory is often called "echoic memory." Attention can transfer incoming sensory information to another short-term memory, known as "working memory." Here information can be manipulated using "phonological" or "scratchpad" stores. Processing the information in this short-term memory through such mechanisms as association and inference enters the information into longer-lasting memories. Auditory memories are important for temporal processing over periods of 1 sec to several seconds. This allows us to discriminate musical rhythms and to follow speech.

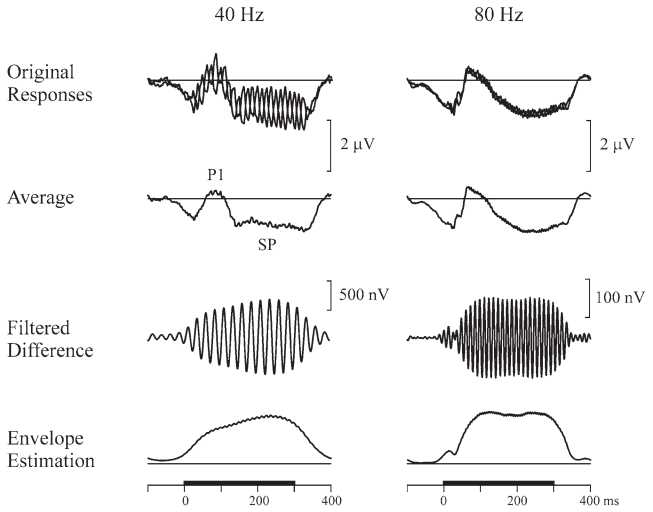

Fig. 4. Temporal integration of the ASSR. In this experiment the stimulus was a continuous 1 kHz tone presented at 65 dB SPL. Every 500 msec there was a 300 msec period of amplitude modulation at either 40 or 80 Hz. The modulation alternated in its onset phase and the responses were separately averaged on the basis of the onset phase. The upper tracings represent the superimposed responses averaged across 10 subjects, with the 80 Hz responses being averaged over four times as many trials as the 40 Hz responses (because the ASSR at the faster rate is much smaller in amplitude). Averaging the responses together (second set of tracings) gives the cerebral response to the modulation. This shows a negative sustained potential through the duration of the modulation and a P1 wave following the onset of the modulation. Because the stimuli were coming at a rate of twice a second there is no recognizable N1 wave. The third line of tracings is the difference between the two original tracings which gives the ASSR with most of the onset response removed. These tracings have been band-pass filtered (35–45 Hz or 70–90 Hz) and amplified with the 80 Hz response amplified more than the 40 Hz response. These waveforms show the rise and fall of the ASSR over the period of the modulation. The lowest set of tracings estimates the envelope of this filtered response. The 40 Hz ASSR develops over approximately 250 msec and ceases within approximately 80 msec after the end of the modulation. The 80 Hz ASSR develops over approximately 100 msec and falls back to baseline within 50 msec after the end of the modulation. Data from Picton (2011, chapter 10). ASSR, auditory steady state responses.

The simplest way to indicate that memory has occurred is to show that the response to a repeated stimulus differs from that to an isolated stimulus. Unlike the earlier responses, the slow cortical AEPs (occurring with latencies greater than 50 msec) have prolonged refractory periods lasting up to 10 or even 20 sec. These slow potentials have therefore been related to sensory and working memories.

Sensory Memory

The N1-P2 response to an auditory stimulus decreases as the interval from a preceding stimulus decreases. This effect can be studied over time by recording the responses to short trains of stimuli and evaluating the decline in response with stimulus repetition, or simply by averaging responses to stimuli presented at different rates (Picton 2011, chapter 11). The rate effect is illustrated in Figure 3, which shows responses recorded with ISIs of 1 and 3 sec. The decrease in amplitude when stimuli are presented more rapidly is often called a refractory period, although whatever is occurring is far more complex than neuronal refractory periods. The time-constants for the rate effect of the N1 wave generated in the temporal lobe are similar to those for auditory sensory memory (Lü et al. 1992; Sams et al. 1993). Sources in different regions of the temporal lobe have different time-constants within the same general range. The rate effects of the N1 wave tend to be nonspecific: they occur even when the stimulus changes. They reflect a stimulus memory that is characterized by only the most basic information: a stimulus has occurred but it is unclear what it is.

More specific memory effects occur with the mismatch negativity (MMN). This small negative wave with a typical latency of approximately 140 msec is superimposed on the response to a deviant sound occurring in a regular sequence of standard sounds (see reviews by Näätänen & Winkler 1999; Picton et al. 2000; Näätänen et al. 2012). The MMN discriminates differences between stimuli with much greater accuracy and along more complex dimensions than the N1. The system underlying the generation of the MMN thus deals with auditory information at a level similar to that of auditory sensory memory as studied behaviorally.

The MMN requires that a standard stimulus be repeated before a stimulus is recognized as deviant: it assesses the regularity as well as the sensory features of the stimuli. A repeating sequence of several stimuli with one deviating from the others should not generate an MMN if the sequence is always the same. Sensory memory should develop a gestalt of the whole sequence and its regular repetition should therefore not be considered deviant. However, this depends on the timing of the sequence. Scherg et al. (1989), presented stimuli once every 0.9 sec and found a clear MMN to a deviant stimulus occurring regularly once every fifth stimulus. This was essentially the same as the MMN when the deviant occurred randomly with

a probability of 0.2. The reason for obtaining an MMN to the regular deviant lies in the span over which the MMN system evaluates sequences. This time period is likely related to the 8 to 10 sec time course of echoic memory. The basic sequence in the Scherg et al. experiment occurred once every 4.5 sec. To detect a regularly repeating sequence, the MMN system would have to maintain in memory at least two full sequences (lasting at least 9 sec) and preferably more. Only if it can evaluate two or more sequences can the system realize that a sequence is repeating itself in a predictable manner. If two sequences lasted 9 sec, a system that had a time span of 8 sec could not determine that there was a regularly repeating sequence and would just react to the deviant tone in the same way as if it occurred randomly. If the stimuli were repeated at a more rapid rate—if the stimulus onset asynchrony was decreased to 100 msec so that the full sequence lasted only 500 msec—no MMN would occur when the sequences were regularly repeated (Sussman et al. 1998). Several sequences would then occur during the MMN time span and the change in frequency would be considered part of a regular sequence and not considered deviant. If the deviant stimulus occurred randomly at the rapid ISI with the same probability as that in the regular presentation (but with no predictable sequence), a clear MMN occurred.

Working Memory

The P3 or P300 component of the AEP is a late positive wave that is typically evoked when an improbable target is actively discriminated from other repeating standard stimuli (reviewed in Picton 2011, chapter 12). This differs from the MMN that occurs whether or not the stimuli are attended. The P300 only occurs when the subject is attending to the stimuli to respond in some way to the occasionally different stimulus. We often indicate this difference by considering this stimulus a "deviant" when it is automatically detected by the MMN system, and a "target" when it is consciously detected by the P300 system. One theory of the P300 is that it represents the updating of context in a working memory that keeps tabs on the different stimuli that are occurring and their relative probabilities (Donchin & Coles 1988). According to this concept the latency of the P300 represents the time taken to evaluate the stimuli before updating working memory.

The amplitude of the P300 wave varies inversely with the probability of the target stimuli. However, this relationship may actually depend more on timing than on probability, because the more improbable target will occur at longer average target-to-target intervals (reviewed by Picton 2011, chapter 12). Gonsalvez and Polich (2002) found that the amplitude of the P300 is critically determined by the target-to-target interval rather than stimulus probability. The P300 amplitude tends to saturate at intervals of over 20 sec. This suggests that information within working memory decays over this period, provided that the information is not re-entered (or "rehearsed").

Changes in the background rhythms of the EEG occur as information is processed in working memory. These changes are not as closely locked in time to the stimulus as the evoked potentials: they are "induced" rather than "evoked." They may be measured by averaging the power or amplitude spectra rather than the time waveform. The rhythms may be increased or decreased during processing: event-related synchronization (ERS) or event-related desynchronization (ERD). The direction of the change will be determined by the level of ongoing activity before the event as well as by the nature of the processing it requires (Pfurtscheller 2006). The discrimination of auditory stimuli causes both a synchronization of θ activity and a desynchronization of α and β activity (Mazaheri & Picton 2005). The encoding of acoustic material into memory elicits α-frequency ERS whereas the retrieval or recognition of the same stimulus material evoked α ERD (reviewed by Krause 2006). ERDs in the α and θ frequencies can be recorded in children in response to auditory stimuli (Fujioka & Ross 2008). These ERDs may be easier to assess during development than the waveform of the late AEP because they do not change as dramatically with increasing age.

Attention and Learning

The amplitude of the P2 wave of the late AEP increases dramatically as subjects learn to discriminate stimuli that they previously found indistinguishable. This finding was initially reported in subjects who learned to discriminate unfamiliar voice-onset times (Tremblay et al. 2001). European languages normally discriminate between voiced and unvoiced stop consonants, such as "b" and "p." For the b phoneme the voicing (vocal cord vibration) occurs at the onset of the sound whereas for the p sound it is delayed by approximately 30 msec. Languages such as Swahili also use an "mb" phoneme wherein the voicing begins before the lips open. English-speaking subjects cannot normally distinguish mb and b but most can be trained to make this discrimination. As they do so the P2 wave increases. Similar results can be obtained in other perceptual learning situations (see reviews by Alain 2007; Tremblay & Moore 2012). However, the late AEP contains multiple overlapping components and each of these can change differently during auditory learning. For example, in the very early stages of learning the P2 wave may decrease rather than increase (Alain et al. 2009).

It remains unclear how much of the P2 effect is specifically related to training and how much related to a nonspecific effect of exposure to the sounds (Sheehan et al. 2005). Perceptual training on a voice-onset time cue changed the AEPs to both the trained stimuli and other untrained stimuli (Ross & Tremblay 2009; Tremblay et al. 2009). Training-specific effects were greater in the left hemisphere, whereas the nonspecific effects were bilateral. The P2 enhancement only occurred in those subjects who improved their ability to discriminate the stimuli. In addition, the subjects who improved showed larger N1 waves (before and after training) than the subjects who did not learn the discrimination. This may be related to the amount of attention paid to the sounds. Experience may nonspecifically enhance responsiveness to many sounds, but attention during the experience may specifically enhance particular categories of sounds.

DISORDERS OF TEMPORAL PROCESSING

Despite the presence of many AEP measurements that can tap into the temporal processing of sounds, and despite the need for objective tests of temporal processing, few AEP tests are currently available for the clinical evaluation of disordered temporal processing. There are several reasons for this state of affairs. Many of the recordings involve more complicated procedures and longer test times than using the ABR to evaluate hearing thresholds or assess brainstem conduction times. Furthermore, signal to noise issues are a particular

problem when responses are small or absent. Does a child with attention deficit disorder have an absent response or is he or she just unable to sit sufficiently still so that small responses can be recognized? Another problem is the lack of normative data: we simply do not know what is within normal limits and what indicates a treatable disorder. In assessing children we also need to know much more about how the AEPs change with development.

Auditory Neuropathy

Starr et al. (1991) described a patient who had difficulty perceiving the temporal aspects of sounds and who had no recognizable ABRs. In a footnote added to the article after it was accepted for publication, the authors reported that the patient showed normal otoacoustic emissions (OAEs). A later article reported on a set of patients with similar findings and coined the term "auditory neuropathy" (Starr et al. 1996). Recent reviews of the disorder have been performed by Starr et al. (2008), Berlin et al. (2010), Picton (2011, chapter 15), and Hood and Morlet (2012). The diagnosis of auditory neuropathy requires two findings, one negative and one positive. These are illustrated in Figure 5.

First, there must be evidence of abnormal processing in the auditory nerve fibers. This is demonstrated by showing that the neural components of the click-ABR are either absent or severely abnormal beginning at wave I. When recording the ABRs, separate responses to condensation and rarefaction clicks must be recorded to distinguish the CM that can sometimes mimic the ABR. The CM completely reverses in polarity for condensation versus rarefaction clicks, whereas the click-ABR shows only slight latency changes.

Second, there must be evidence for preserved function in the outer hair cells. This is most clearly demonstrated by recording OAEs. In approximately a quarter of the patients with auditory neuropathy, OAEs are absent. In these cases, preservation of function in the outer hair cells can be demonstrated by recording the CM. This must be clearly differentiated from electrical stimulus artifact by connecting the transducer to a tube leading to an ear-insert. The tube causes an acoustic delay, which makes the stimulus artifact occur significantly earlier than the CM.

AEPs occurring later than the ABR may be normal or abnormal in patients with auditory neuropathy, depending in part on the severity of the disorder. The N1 waves evoked by brief tones are significantly delayed in patients with auditory neuropathy (Michalewski et al. 2009). The N1-P2 response to gaps in continuous noise can be recognized in patients with auditory neuropathy only when the gaps exceed 10 msec, whereas normal subjects show clear responses to 5 msec gaps (Michalewski et al. 2005). The N1 wave to a change in the intensity or frequency of an ongoing tone is delayed or absent in patients with auditory neuropathy, particularly in patients with the postsynaptic form of the disorder (Dimitrijevic et al. 2011).

Auditory neuropathy can be caused in many different ways. The disorder presents at two different ages: in infancy, patients are referred when abnormal ABRs are found during newborn hearing screening; in later years patients present because they are experiencing hearing difficulties. In infancy, the main etiological factors are genetic disorders, hyperbilirubinemia, and hypoxia. Older patients presenting with auditory neuropathy may have hereditary sensorimotor neuropathy, acquired neuropathies, infections, or toxicity.

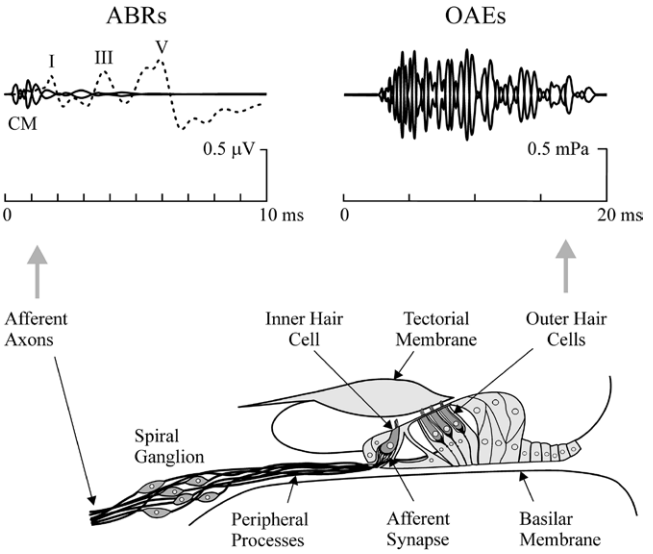

Fig. 5. Auditory neuropathy. The upper part shows the typical physiological responses of a patient with auditory neuropathy. Responses to condensation and rarefaction clicks are superimposed. On the left are the ABRs that derive from activation of the auditory nerve axons and the brainstem pathways. The patient's responses are represented using continuous lines and the normal pattern of response is shown by the dotted lines. The patient shows no recognizable ABR but a small cochlear microphonic that reverses polarity with the stimuli. On the left are shown the otoacoustic emissions that derive mainly from the outer hair cells. These show a normal pattern. The lower half shows the organ of Corti and the afferent nerve fibers. Auditory neuropathy may be caused by a disorder of the tectorial membrane, the inner hair cell, the afferent synapses between the inner hair cell and the peripheral processes of the auditory neurons in the spiral ganglion, or the afferent axons. ABR, auditory brainstem responses.

Auditory neuropathy is a disorder of the afferent auditory nervous system: the inner hair cells and the afferent nerve fibers of the auditory nerve. The pathophysiology of the disorder may be presynaptic (involving the inner hair cell or its synapse) or postsynaptic (involving the afferent neurons). Electrocochleography may help to distinguish these different abnormalities (McMahon et al. 2008; Santarelli et al. 2008). The disorders of the auditory nerve may be further divided into those that involve the nerve fibers or the myelin sheath. The absence or distortion of the ABR may be caused by either a loss of neurons (depletion) or by abnormalities of their conduction velocity (desynchrony).

Auditory neuropathy manifests as a disorder of auditory temporal processing. Hearing thresholds may be normal. Even when the thresholds are elevated, the ability to process speech is reduced far more than what could be expected from the changes in threshold. Patients have much more difficulty making time-based rather than frequency-based discriminations (Zeng et al. 2005).

Aging

Disorders of auditory temporal processing are common in the elderly. As we grow older we have difficulty detecting gaps in sounds (Schneider & Hamstra 1999; Humes et al. 2010), discriminating modulated from unmodulated sounds (He et al. 2008), and recognizing mistuned harmonics (Alain et al. 2001a; Alain & McDonald 2007).

As previously discussed, clear N1-P2 responses can be recorded to changes in the interaural phase of low-frequency binaural sounds, provided the change occurs at the null point of an ongoing amplitude modulation. Magnetic responses to this stimulus have been recorded in young, middle-aged, and elderly subjects (Ross et al. 2007a and b). The response can be recognized with carrier frequencies up to 1250 Hz in young subjects, but in elderly subjects, it can only be recorded up to 760 Hz. Middle-aged subjects showed no recognizable response to phase changes for carrier frequencies higher than 940 Hz, indicating that our binaural timing abilities begin to decline in mid life.

Ross et al. (2010) found that the response of the human auditory cortex to a gap between two brief noise bursts decreased in amplitude with increasing age. These physiological changes were related to a decreased perceptual ability to detect the gap. Both the physiological and the perceptual changes began in middle age. Similarly, the AEPs to occasional gaps in an ongoing noise are markedly decreased as we grow older (Harris et al. 2012). Alain et al. (2004) showed that the elderly can still discriminate occasional brief stimuli with gaps from those without gaps (and generate the P3 waves associated with the discrimination) even when the stimuli do not elicit any recognizable MMN when unattended. The authors attributed these results to top–down compensation for age-related deficits in automatic stimulus evaluation.

Aging significantly affects human following responses. The FFR decreases significantly in amplitude with increasing age, although these changes occur independently of changes in our ability to discriminate frequencies (Clinard et al. 2010). The electrophysiological response to binaural beats decreases in the elderly (Grose & Mamo 2012). Changes also occur in the envelope following response as we grow older. The frequency at which the 40 Hz response has its maximum amplitude decreases, the apparent latency of the response increases and the response at modulation rates greater than 80 Hz decreases in amplitude (Purcell et al. 2004; Picton 2011).

The MMN is decreased in the elderly. This may be related to deterioration in auditory sensory memory, or to decreased signal to noise ratio of the information reaching sensory memory (Alain et al. 2004). Unfortunately, the MMN is abnormal in many different disorders (see Näätänen et al. 2012 for a recent review) and has little diagnostic specificity. An abnormal MMN indicates that all is not well but does not tell us much about why.

Dyslexia

Many children experience difficulty learning to read (Snowling & Göbel 2011). Approximately 5% of children have a disorder known as "developmental dyslexia." This disorder often runs in families, but little is known about its pathophysiology. Diagnosis is readily made on the basis of the level of reading achievement compared with that in normal children of the same age. However, reading disability may itself be part of other disorders such as learning disability, attention deficit hyperactivity disorder, auditory processing disorder, speech sound impairment, and language impairment (Dawes & Bishop 2009; Pennington & Bishop 2009). In addition, reading disorders will necessarily vary across languages: the neural processing needed for reading differs between alphabetic and ideographic orthographies and between languages with consistent or inconsistent relations between sound and symbol (Hadzibeganovic et al. 2010).

Many scientists have considered the possibility that dyslexia might be caused by a simple sensory disorder. The two most commonly proposed abnormalities involve the magnocellular visual system, which controls ocular movements and fixation (e.g., Stein 2001) and the temporal processing of auditory information (e.g., Tallal 1980). Other scientists have proposed that these specific abnormalities might themselves be part of a more general brain disorder with variable manifestations (e.g., Galaburda et al. 2006). Speech processing difficulties occur in dyslexic subjects more than in normal readers, but the abnormalities are not consistent across subjects and do not correlate with the severity of the dyslexia (Georgiou et al. 2010; Messaoud-Galusi et al. 2011).

The auditory temporal processing theory of dyslexia proposes a deficit in forming accurate neural representations for speech sounds caused by an impaired of ability to follow rapid auditory changes. Ill-defined speech representations would then make it difficult to relate speech sounds to alphabetic codes. If so, children with dyslexia could be helped by training them to recognize rapid auditory changes. This idea led to a series of training programs called Fast ForWord. Although early reports suggested that these programs significantly improved reading ability (Merzenich et al. 1996; Tallal 2004), more recent studies have found no clear benefit (Borman et al. 2009; Strong et al. 2011) or benefits not significantly different from those of other therapies (Loeb et al. 2009).

Nevertheless, problems with auditory temporal processing remain common in dyslexic individuals. Dyslexic subjects do not detect amplitude modulation at rates greater than 10 Hz as well as normal subjects do and show smaller ASSRs to these stimuli (Menell et al. 1999; Poelmans et al. 2012). Vandermosten et al. (2010) showed that dyslexic subjects did not discriminate rapidly changing speech sounds (such as stop consonants) as accurately as equally difficult speech sounds differing in spectra rather than timing (such as vowels). These problems persisted when the spectra of the stimuli were altered so that they were not perceived as speech. The deficit therefore involves temporal processing and is not specific to speech.

Changes in the overall speech envelope distinguish different syllables, and following speech envelopes is essential to recognizing changes from one phoneme to another. A disorder in following the speech envelope could therefore disrupt the formation of stable phonological representations. Goswami et al. (2002) has suggested that the disorder might be related to detecting the changes in the speech envelope that signal syllable transitions. Such a detection mechanism might involve the slow electrical rhythms in the cortex, which are entrained by changing phonological information (Goswami 2011).

Recent physiological studies have found that dyslexic subjects show reduced following responses at slow modulation rates. Abrams et al. (2009) found abnormalities in following the speech envelope at syllabic rates. Poelmans et al. (2012) found abnormalities in the steady state responses to modulated speech noise at phonemic rates (20 Hz). Hämäläinenet al. (2012) found abnormalities at very slow rates (2 Hz).

Auditory Processing Disorders

Deficits in discriminating sounds, especially on the basis of their timing, can occur in the absence of raised auditory thresholds. These problems are often attributed to a "central auditory processing disorder" (CAPD). However, this is not a clear

diagnostic entity and there is great overlap with other clinical entities such as learning disorder, attention deficit disorder, and specific language impairment (Cacace & McFarland 2009; Dawes & Bishop 2009; Moore et al. 2010). The diagnosis of children with these developmental problems often depends on who is doing the evaluation and what tests are performed.

Probably, there are many different kinds of CAPD. Some may affect simple brainstem mechanisms such as those used to localize a particular speaker and thus differentiate that person's speech from others. Others may affect higher levels of perception, such as the ability to follow rapid phonemic changes. Some may be specific to the auditory system whereas others may involve supramodal language impairments that affect reading as well as hearing.

One problem with assessing these disorders is the interaction with development. We become better at discriminating the temporal aspects of auditory stimuli as we grow older. Part of this improvement may depend on learning and experience, and part may just depend on the maturation of the nervous system. If a child is unable to process the temporal aspects of speech during the acquisition of auditory language, language processing may not develop properly even if the temporal processing later reaches normal levels. A causal relation between temporal processing and language performance "might be most important or evident during language acquisition or while learning to read, and not necessarily after language and/or reading performance has reached any adult plateau" (Phillips 2012, p. 93). If this is so, we may need to diagnose the defect and initiate therapy before the symptoms of the disorder become apparent.

Perception involves both top–down and bottom–up processing. We perceive auditory stimuli better when we know what to expect. If we can model what we perceive (because we have experience perceiving it and because we are focusing our attention on it), we perform better than if we just passively listen. We must therefore always consider cognition when assessing disorders of speech and language. Across-channel gap detection, for example, varies much more from subject to subject than within-channel gap detection, and this variability may depend on cognitive rather than auditory factors (Phillips 2012). Training programs to help children with listening problems must therefore develop general cognitive abilities for attention and memory at the same time as improving basic auditory skills (Moore et al. 2009).

AEPs might be able to distinguish the different pathophysiologies that lead to CAPD. For example, a patient may demonstrate abnormalities in some AEP components but not in others. A patient with auditory neuropathy may be indistinguishable from a patient with CAPD unless the ABR is recorded. Perhaps an electrophysiological test battery can be constructed that will discriminate among different causes of CAPD. Such a battery should include the ABR , the FFR, and the speech ABR of Johnson et al. (2005). Additional measurements that would be helpful include the cortical responses to binaural processing, the following responses entrained by the speech envelope, the late AEPs evoked by gaps in a continuous noise, and the ERD that accompanies meaningful auditory processing. Some early studies are promising. When compared with normally developing children, children with language impairments show smaller FFRs and delayed ABRs at rapid rates (Basu et al. 2010), abnormal speech ABRs and an increased sensitivity of the late AEPs to noise (Banai et al. 2009; Hornickel & Kraus 2011; Wible et al. 2005), abnormal cortical responses to the speech envelope (Abrams et al. 2009), and decreased low-frequency ERDs to auditory stimuli (Bishop et al. 2010).

CONCLUSION

Auditory neurons follow the timing of the sounds we hear. They can recognize interaural difference of some tens of microseconds. They can accurately portray the frequency of a sound up to approximately 1500 Hz. They can track the envelopes of ongoing speech to give us the pitch of the speaker and the changes between the phonemes. At very slow rates they may lock our thoughts to the prosody of language. Neuronal networks integrate information over the periods necessary to process the location of a sound and to determine what it means. They maintain the information in simple form for periods of several seconds and in meaningful form for periods of several tens of seconds. Ultimately they fix important sounds in long-term memory so that we can later perceive and interpret sounds more easily.

The basic ideas of auditory temporal processing can be seen in the clock diagram of Figure 6. This "clock to tell the time of hearing" probably belongs to the White Rabbit from Lewis Carroll's *Alice's Adventures in Wonderland*. At least, his portrait by John Tenniel graces the clock face. Our clock considers time differently from ordinary clocks. The scale—from 0.10 msec to 100 sec—is logarithmic rather than linear. The small hand points to the timing of what is perceived. These phenomena are listed outside the edge of the clock according to their modal times. For example, typically recognizable gaps last approximately 2 msec, phonemes approximately 100 msec (consonants can be as short as several tens of milliseconds and vowels may last several hundreds), and echoic memory several seconds. The large hand points to what is happening in the brain and recorded in the AEPs. This clock works differently from an ordinary clock in that the small hand drives the large one rather than vice versa. As the small hand steps through each of the different auditory phenomena it initiates a whole sequence of auditory processes

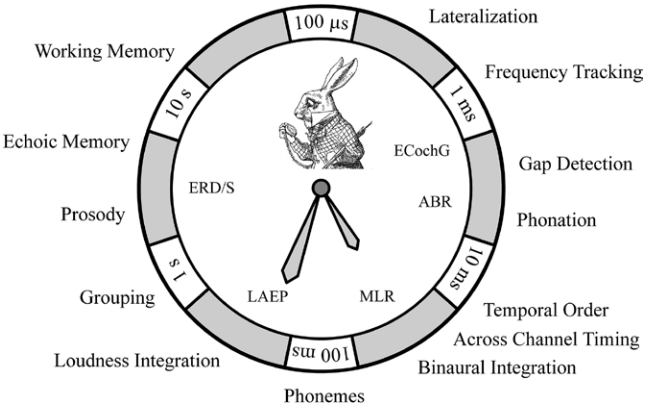

Fig. 6. A Clock to tell the time of hearing. The timings for different auditory stimuli are listed around the outer edge of the clock. The AEPs are listed within the clock face. These go from the electrocochlegram in the first few milliseconds after the stimulus to the periods of event-related synchronization and desynchronization which can last for several seconds. The small hand moves through the different stimuli. Each stimulus triggers the large hand to proceed through all the physiological responses it evokes. The hands are set to indicate the late AEP to a change in the interaural time differences (as shown in Fig.2). AEP, auditory evoked potential.

in the human ear and brain. These are indicated by the revolution of the large hand as the brain generates the AEPs listed within the edge of the clock. As each auditory phenomenon occurs, it is registered in the cochlea, transmitted through the brainstem, projected to the cortex, and ultimately sent to memory. The particular example indicated by the clock hands in the figure is the late AEP to a change in binaural timing (as shown in Fig. 2).

Disorders of temporal processing occur in many different clinical syndromes. Auditory neuropathy is a disorder of the peripheral afferent nervous system (inner hair cell and the afferent nerve fibers). Its primary manifestation is decreased perception of auditory timing information. As we age we tend to have decreased hearing. Some of this is related to changes in the cochlea and is shown by elevated thresholds, but some is also related to the aging nervous system's difficulty in following rapidly changing sounds. Children with developmental disorders such as dyslexia or CAPD commonly show abnormalities of auditory temporal processing. However, the nature of these abnormalities and how they might contribute to the cognitive disorder remains to be determined. A battery of electrophysiological tests might be helpful in categorizing different disorders and in monitoring treatment. This will require extensive normative data.

ACKNOWLEDGMENT

The author receives remuneration from Baycrest based on sales of the Master system by Natus Medical Incorporated, and royalties from Plural Publishing Incorporated for the book *Human Auditory Evoked Potentials*.

The author declares no other conflict of interest.

Address for correspondence: Terence W. Picton, 79 Burnett Ave, Toronto, Ontario, Canada M2N 1V4. E-mail: terry.picton@gmail.com

Received September 10, 2012; accepted October 21, 2012.

REFERENCES

Abrams, D. A., Nicol, T., Zecker, S., et al. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J Neurosci*, *28*, 3958–3965.

Abrams, D. A., Nicol, T., Zecker, S., et al. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *J Neurosci*, *29*, 7686–7693.

Ahissar, E., Nagarajan, S., Ahissar, M., et al. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci USA*, *98*, 13367–13372.

Aiken, S. J., & Picton, T. W. (2006). Envelope following responses to natural vowels. *Audiol Neurootol*, *11*, 213–232.

Aiken, S. J., & Picton, T. W. (2008a). Envelope and spectral frequency-following responses to vowel sounds. *Hear Res*, *245*, 35–47.

Aiken, S. J., & Picton, T. W. (2008b). Human cortical responses to the speech envelope. *Ear Hear*, *29*, 139–157.

Alaerts, J., Luts, H., Hofmann, M., et al. (2009). Cortical auditory steady-state responses to low modulation rates. *Int J Audiol*, *48*, 582–593.

Alain, C. (2007). Breaking the wave: Effects of attention and learning on concurrent sound perception. *Hear Res*, *229*, 225–236.

Alain, C., Campeanu, S., Tremblay, K. (2010). Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *J Cogn Neurosci*, *22*, 392–403.

Alain, C., McDonald, K. L., Ostroff, J. M., et al. (2001a). Age-related changes in detecting a mistuned harmonic. *J Acoust Soc Am*, *109*, 2211–2216.

Alain, C., McDonald, K. L., Ostroff, J. M., et al. (2004). Aging: A switch from automatic to controlled processing of sounds? *Psychol Aging*, *19*, 125–133.

Alain, C., Arnott, S. R., Picton, T. W. (2001b). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *J Exp Psychol Hum Percept Perform*, *27*, 1072–1089.

Alain, C., & McDonald, K. L. (2007). Age-related differences in neuromagnetic brain activity underlying concurrent sound perception. *J Neurosci*, *27*, 1308–1314.

Alain, C., Reinke, K., He, Y., et al. (2005). Hearing two things at once: Neurophysiological indices of speech segregation and identification. *J Cogn Neurosci*, *17*, 811–818.

Alain, C., & Winkler, I. (2012). Recording event-related brain potentials: Application to study auditory perception. In D. Poeppel, T. Overath, A. N. Popper, et al. (Eds). *The Human Auditory Cortex. Springer Handbook of Auditory Research* (*Vol. 43*, pp. 69–96). New York, NY: Springer-Verlag.

Banai, K., Hornickel, J., Skoe, E., et al. (2009). Reading and subcortical auditory function. *Cereb Cortex*, *19*, 2699–2707.

Basu, M., Krishnan, A., Weber-Fox, C. (2010). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Dev Sci*, *13*, 77–91.

Berlin, C. I., Hood, L. J., Morlet, T., et al. (2010). Multi-site diagnosis and management of 260 patients with auditory neuropathy/dys-synchrony (auditory neuropathy spectrum disorder). *Int J Audiol*, *49*, 30–43.

Bidelman, G. M., Gandour, J. T., Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci*, *23*, 425–434.

Bishop, D. V., Hardiman, M. J., Barry, J. G. (2010). Lower-frequency event-related desynchronization: A signature of late mismatch responses to sounds, which is reduced or absent in children with specific language impairment. *J Neurosci*, *30*, 15578–15584.

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

Borman, G. D., Benson, J. G., & Overman, L. (2009). A randomized field trial of the Fast ForWord language computer-based training program. *Educational Evaluation and Policy Analysis*, *31*, 82–106.

Bourguignon, M., De Tiège, X., de Beeck, M. O., et al. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Hum Brain Mapp*, 34, 314–326.

Butler, B. E., & Trainor, L. J. (2012). Sequencing the cortical processing of pitch-evoking stimuli using EEG analysis and source estimation. *Front Psychol*, *3*, 180.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: Bradford Books.

Brugge, J. F., Nourski, K. V., Oya, H., et al. (2009). Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J Neurophysiol*, *102*, 2358–2374.

Budd, T. W., & Michie, P. T. (1994). Facilitation of the N1 peak of the auditory ERP at short stimulus intervals. *Neuroreport*, *5*, 2513–2516.

Cacace, A. T., & McFarland, D. J. (Eds). (2009). *Controversies in Central Auditory Processing Disorder.* San Diego, CA: Plural Publishing

Clinard, C. G., Tremblay, K. L., Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hear Res*, *264*, 48–55.

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

Dajani, H. R., & Picton, T. W. (2006). Human auditory steady-state responses to changes in interaural correlation. *Hear Res*, *219*, 85–100.

Damaschke, J., Riedel, H., Kollmeier, B. (2005). Neural correlates of the precedence effect in auditory evoked potentials. *Hear Res*, *205*, 157–171.

Dawes, P., & Bishop, D. (2009). Auditory processing disorder in relation to developmental disorders of language, communication and attention: A review and critique. *Int J Lang Commun Disord*, *44*, 440–465.

Dimitrijevic, A., Starr, A., Bhatt, S., et al. (2011). Auditory cortical N100 in pre- and post-synaptic auditory neuropathy to frequency or intensity changes of continuous tones. *Clin Neurophysiol*, *122*, 594–604.

Dingle, R. N., Hall, S. E., Phillips, D. P. (2010). A midline azimuthal channel in human spatial hearing. *Hear Res*, *268*, 67–74.

Divenyi, P. L. (2004). The times of Ira Hirsh Multiple ranges of auditory temporal perception. *Semin Hear*, *25*, 229–239.

Dobie, R. A., & Norton, S. J. (1980). Binaural interaction in human auditory evoked potentials. *Electroencephalogr Clin Neurophysiol*, *49*, 303–313.

Donchin, E., Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behav Brain Sci*, *11*, 357–374.

Eddins, D. A., & Green, D. A. (1995). Temporal integration and temporal resolution. In B. C. J. Moore (Ed). *Hearing: Handbook of Perception and Cognition* (2nd ed., pp. 207–242). San Diego, CA: Academic Press.

Elangovan, S., & Stuart, A. (2008). Natural boundaries in gap detection are related to categorical perception of stop consonants. *Ear Hear*, *29*, 761–774.

Fox, A. M., Anderson, M., Reid, C., et al. (2010). Maturation of auditory temporal integration and inhibition assessed with event-related potentials (ERPs). *BMC Neurosci*, *11*, 49.

Fox, A. M., Reid, C. L., Anderson, M., et al. (2012). Maturation of rapid auditory temporal processing and subsequent nonword repetition performance in children. *Dev Sci*, *15*, 204–211.

Friesen, L. M., & Picton, T. W. (2010). A method for removing cochlear implant artifact. *Hear Res*, *259*, 95–106.

Fujioka, T., & Ross, B. (2008). Auditory processing indexed by stimulus-induced alpha desynchronization in children. *Int J Psychophysiol*, *68*, 130–140.

Galaburda, A. M., LoTurco, J., Ramus, F., et al. (2006). From genes to behavior in developmental dyslexia. *Nat Neurosci*, *9*, 1213–1217.

Galbraith, G. C., Bhuta, S. M., Choate, A. K., et al. (1998). Brain stem frequency-following response to dichotic vowels during attention. *Neuroreport*, *9*, 1889–1893.

Georgiou, G. K., Protopapas, A., Papadopoulos, T. C., et al. (2010). Auditory temporal processing and dyslexia in an orthographically consistent language. *Cortex*, *46*, 1330–1344.

Gockel, H. E., Carlyon, R. P., Mehta, A., et al. (2011). The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *J Assoc Res Otolaryngol*, *12*, 767–782.

Gonsalvez, C. J., & Polich, J. (2002). The target-to-target interval is the critical determinant of the P3. *Psychophysiology*, *39*, 388–396.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn Sci*, *15*, 3–10.

Goswami, U., Thomson, J., Richardson, U., et al. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proc Natl Acad Sci USA*, *99*, 10911–10916.

Grantham, D. W. (1995). Spatial hearing and related phenomena. In B. C. J. Moore (Ed). *Hearing: Handbook of Perception and Cognition* (2nd ed., pp. 297–345). San Diego, CA: Academic Press.

Grantham, D. W., & Wightman, F. L. (1978). Detectability of varying interaural temporal differences. *J Acoust Soc Am*, *63*, 511–523.

Grose, J. H., & Mamo, S. K. (2012). Electrophysiological measurement of binaural beats: Effects of primary tone frequency and observer age. *Ear Hear*, *33*, 187–194.

Gutschalk, A., Micheyl, C., Melcher, J. R., et al. (2005). Neuromagnetic correlates of streaming in human auditory cortex. *J Neurosci*, *25*, 5382–5388.

Hadzibeganovic, T., van den Noort, M., Bosch, P., et al. (2010). Cross-linguistic neuroimaging and dyslexia: A critical view. *Cortex*, *46*, 1312–1316.

Halliday, R., & Callaway, E. (1978). Time shift evoked potentials (TSEPs): Method and basic results. *Electroencephalogr Clin Neurophysiol*, *45*, 118–121.

Hämäläinen, J. A., Rupp, A., Soltész, F., et al. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: An MEG study. *Neuroimage*, *59*, 2952–2961.

Harris, K. C., Wilson, S., Eckert, M. A., et al. (2012). Human evoked cortical activity to silent gaps in noise: Effects of age, attention, and cortical processing speed. *Ear Hear*, *33*, 330–339.

Hartmann, W. M., McAdams, S., Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *J Acoust Soc Am*, *88*, 1712–1724.

He, N. J., Mills, J. H., Ahlstrom, J. B., et al. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *J Acoust Soc Am*, *124*, 3841–3849.

Hertrich, I., Dietrich, S., Trouvain, J., et al. (2012). Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal. *Psychophysiology*, *49*, 322–334.

Hertrich, I., Mathiak, K., Lutzenberger, W., et al. (2004). Time course and hemispheric lateralization effects of complex pitch processing: Evoked magnetic fields in response to rippled noise stimuli. *Neuropsychologia*, *42*, 1814–1826.

Hirsh, I. J. (1959). Auditory perception of temporal order. *J Acoust Soc Am*, *31*,759–767.

Howard, M. F., & Poeppel, D. (2012). The neuromagnetic response to spoken sentences: Co-modulation of theta band amplitude and phase. *Neuroimage*, *60*, 2118–2127.

Hood, L. J., & Morlet, T. (2012). Current issues in auditory neuropathy spectrum disorder. In K. Tremblay & R. Burkard (Eds). *Translational Perspectives in Auditory Neuroscience: Special Topics* (pp. 35–68). San Diego, CA: Plural Publishing.

Hornickel, J., & Kraus, N. (2011). Objective biological measures for the assessment and management of Auditory Processing Disorder. *Curr Pediatr Rev*, 7, 252–261.

Humes, L. E., Kewley-Port, D., Fogerty, D., et al. (2010). Measures of hearing threshold and temporal processing across the adult lifespan. *Hear Res*, *264*, 30–40.

Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci USA*, *101*, 6809–6814.

Johnson, K. L., Nicol, T. G., Kraus, N. (2005). Brain stem response to speech: A biological marker of auditory processing. *Ear Hear*, *26*, 424–434.

Jones, S. J., Pitman, J. R., Halliday, A. M. (1991). Scalp potentials following sudden coherence and discoherence of binaural noise and change in the inter-aural time difference: A specific binaural evoked potential or a "mismatch" response? *Electroencephalogr Clin Neurophysiol*, *80*, 146–154.

Joris, P. X., van de Sande, B., Recio-Spinoso, A., et al. (2006). Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. *J Neurosci*, *26*, 279–289.

Koskinen, M., Viinikanoja, J., Kurimo, M., et al. (2012). Identifying fragments of natural speech from the listeners's MEG signals. *Hum Brain Mapp* [in press. Published online 17 Feb 2012], DOI: 10.1002/hbm.22004.

Krause, C. M. (2006). Cognition- and memory-related ERD/ERS responses in the auditory stimulus modality. *Prog Brain Res*, *159*, 197–207.

Krishnan, A. (2007) Frequency-following response. In R. F. Burkard, M. Don, & J. J. Eggermont (Eds). *Auditory Evoked Potentials: Basic Principles and Clinical Applications* (pp. 313–333). Baltimore, MD: Lippincott, Williams & Wilkins.

Krishnan, A., Xu, Y., Gandour, J. T., et al. (2004). Human frequency-following response: Representation of pitch contours in Chinese tones. *Hear Res*, *189*, 1–12.

Krishnan, A., Xu, Y., Gandour, J., et al. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res*, *25*, 161–168.

Krumbholz, K., Patterson, R. D., Seither-Preisler, A., et al. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb Cortex*, *13*, 765–772.

Lalor, E. C., & Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur J Neurosci*, *31*, 189–193.

Li, L., Qi, J. G., He, Y., et al. (2005). Attribute capture in the precedence effect for long-duration noise sounds. *Hear Res*, *202*, 235–247.

Liebenthal, E., & Pratt, H. (1999). Human auditory cortex electrophysiological correlates of the precedence effect: Binaural echo lateralization suppression. *J Acoust Soc Am*, *106*, 291–303.

Loeb, D. F., Gillam, R. B., Hoffman, L., et al. (2009). The effects of Fast ForWord Language on the phonemic awareness and reading skills of school-age children with language impairments and poor reading skills. *Am J Speech Lang Pathol*, *18*, 376–387.

Loveless, N., Levänen, S., Jousmäki, V., et al. (1996). Temporal integration in auditory sensory memory: Neuromagnetic evidence. *Electroencephalogr Clin Neurophysiol*, *100*, 220–228.

Lü, Z. L., Williamson, S. J., Kaufman, L. (1992). Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Res*, *572*, 236–241.

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

Mazaheri, A., & Picton, T. W. (2005). EEG spectral dynamics during discrimination of auditory and visual targets. *Brain Res Cogn Brain Res*, *24*, 81–96.

McEvoy, L., Levänen, S., Loveless, N. (1997). Temporal characteristics of auditory sensory memory: Neuromagnetic evidence. *Psychophysiology*, *34*, 308–316.

McEvoy, L. K., Picton, T. W., Champagne, S. C., et al. (1990). Human evoked potentials to shifts in the lateralization of a noise. *Audiology*, *29*, 163–180.

McEvoy, L. K., Picton, T. W., Champagne, S. C. (1991). Effects of stimulus parameters on human evoked potentials to shifts in the lateralization of a noise. *Audiology*, *30*, 286–302.

McMahon, C. M., Patuzzi, R. B., Gibson, W. P., et al. (2008). Frequency-specific electrocochleography indicates that presynaptic and postsynaptic mechanisms of auditory neuropathy exist. *Ear Hear*, *29*, 314–325.

McPherson, D. L., & Starr, A. (1993). Binaural interaction in auditory evoked potentials: Brainstem, middle- and long-latency components. *Hear Res*, *66*, 91–98.

McPherson, D. L., & Starr, A. (1995). Auditory time-intensity cues in the binaural interaction component of the auditory evoked potentials. *Hear Res*, *89*, 162–171.

Menell, P., McAnally, K. I., Stein, J. F. (1999). Psychophysical sensitivity and physiological response to amplitude modulation in adult dyslexic listeners. *J Speech Lang Hear Res*, *42*, 797–803.

Merzenich, M. M., Jenkins, W. M., Johnston, P., et al. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, *271*, 77–81.

Messaoud-Galusi, S., Hazan, V., Rosen, S. (2011). Investigating speech perception in children with dyslexia: Is there evidence of a consistent deficit in individuals? *J Speech Lang Hear Res*, *54*, 1682–1701.

Michalewski, H. J., Starr, A., Nguyen, T. T., et al. (2005). Auditory temporal processes in normal-hearing individuals and in patients with auditory neuropathy. *Clin Neurophysiol*, *116*, 669–680.

Michalewski, H. J., Starr, A., Zeng, F. G., et al. (2009). N100 cortical potentials accompanying disrupted auditory nerve activity in auditory neuropathy (AN): Effects of signal intensity and continuous noise. *Clin Neurophysiol*, *120*, 1352–1363.

Mills, A. W. (1958). On the minimum audible angle, *J Acoust Soc Am*, *30*, 237–246.

Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing* (5th ed.). Amsterdam, The Netherlands: Academic Press (Elsevier Science).

Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., et al. (2010). Nature of auditory processing disorder in children. *Pediatrics*, *126*, e382–e390.

Moore, D. R., Halliday, L. F., Amitay, S. (2009). Use of auditory learning to manage listening problems in children. *Philos Trans R Soc Lond, B, Biol Sci*, *364*, 409–420.

Näätänen, R., Kujala, T., Escera, C., et al. (2012). The mismatch negativity (MMN)–A unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clin Neurophysiol*, *123*, 424–458.

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull*, *125*, 826–859.

Nourski, K. V., Reale, R. A., Oya, H., et al. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *J Neurosci*, *29*, 15564–15574.

Ozdamar, O., Bohorquez, J., Mihajloski, T., et al. (2011). Auditory evoked responses to binaural beat illusion: Stimulus generation and the derivation of the binaural interaction component (BIC). *Conf Proc IEEE Eng Med Biol Soc*, *2011*, 830–833.

Pasley, B. N., David, S. V., Mesgarani, N., et al. (2012). Reconstructing speech from human auditory cortex. *PLoS Biol*, *10*, e1001251.

Pennington, B. F., & Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annu Rev Psychol*, *60*, 283–306.

Pfurtscheller, G. (2006). The cortical activation model (CAM). *Prog Brain Res*, *159*, 19–27.

Phillips, D. P. (2012). Time and timing in audition: Some current issues in auditory temporal processing. In K. Tremblay & R. Burkard (Eds). *Translational Perspectives in Auditory Neuroscience: Special Topics* (pp. 69–101). San Diego, CA: Plural Publishing.

Phillips, D. P., & Smith, J. C. (2004). Correlations among within-channel and between-channel auditory gap-detection thresholds in normal listeners. *Perception*, *33*, 371–378.

Picton T. W. (2011). *Human Auditory Evoked Potentials.* San Diego, CA: Plural Press.

Picton, T. W., Alain, C., Otten, L., et al. (2000). Mismatch negativity: Different water in the same river. *Audiol Neurootol*, *5*, 111–139.

Picton, T. W., & Ross, B. (2010). Physiological measurements of human binaural processing. In J. Buchholz., T. Dau, J. C. Dalsgaard, et al. (Eds). *Binaural processing and spatial hearing: Proceedings of the Second International Symposium on Audiological and Auditory Research (ISAAR 2009)* (pp. 15–28). Helsingør, Denmark: Danavox Jubilee Foundation.

Picton, T. W., Skinner, C. R., Champagne, S. C., et al. (1987). Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *J Acoust Soc Am*, *82*, 165–178.

Picton, T. W., & Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Dev Neuropsychol*, *31*, 249–278.

Poelmans, H., Luts, H., Vandermosten, M., et al. (2012). Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia. *Ear Hear*, *33*, 134–143.

Poth, E. A., Boettcher, F. A., Mills, J. H., et al. (2001). Auditory brainstem responses in younger and older adults for broadband noises separated by a silent gap. *Hear Res*, *161*, 81–86.

Pratt, H., Starr, A., Michalewski, H. J., et al. (2010). A comparison of auditory evoked potentials to acoustic beats and to binaural beats. *Hear Res*, *262*, 34–44.

Purcell, D. W., John, S. M., Schneider, B. A., et al. (2004). Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am*, *116*, 3581–3593.

Rasch, R., & Plomp, R. 1999. The perception of musical tones. In D. Deutsch (Ed). *The Psychology of Music* (2nd ed., pp. 89–112). New York, NY: Academic Press.

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

Riedel, H., & Kollmeier, B. (2006). Interaural delay-dependent changes in the binaural difference potential of the human auditory brain stem response. *Hear Res*, *218*, 5–19.

Ross, B. (2008). A novel type of auditory responses: Temporal dynamics of 40-Hz steady-state responses induced by changes in sound localization. *J Neurophysiol*, *100*, 1265–1277.

Ross, B., Fujioka, T., Tremblay, K. L., et al. (2007a). Aging in binaural hearing begins in mid-life: Evidence from cortical auditory-evoked responses to changes in interaural phase. *J Neurosci*, *27*, 11172–11178.

Ross, B., Herdman, A. T., Pantev, C. (2005). Stimulus induced desynchronization of human auditory 40-Hz steady-state responses. *J Neurophysiol*, *94*, 4082–4093.

Ross, B., Picton, T. W., Pantev, C. (2002). Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hear Res*, *165*, 68–84.

Ross, B., Schneider, B., Snyder, J. S., et al. (2010). Biological markers of auditory gap detection in young, middle-aged, and older adults. *PLoS ONE*, *5*, e10101.

Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear Res*, *248*, 48–59.

Ross, B., Tremblay, K. L., Picton, T. W. (2007b). Physiological detection of interaural phase differences. *J Acoust Soc Am*, *121*, 1017–1027.

Sable, J. J., Low, K. A., Maclin, E. L., et al. (2004). Latent inhibition mediates N1 attenuation to repeating sounds. *Psychophysiology*, *41*, 636–642.

Sams, M., Hari, R., Rif, J., et al. (1993). The human auditory sensory memory trace persists about 10 s: Neuromagnetic evidence. *J Cog Neurosci*, *5*, 363–370.

Santarelli, R., Starr, A., Michalewski, H. J., et al. (2008). Neural and receptor cochlear potentials obtained by transtympanic electrocochleography in auditory neuropathy. *Clin Neurophysiol*, *119*, 1028–1041.

Scherg, M., Vajsar, J., Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *J Cog Neurosci*, *1*, 336–355.

Schneider, B. A., & Hamstra, S. J. (1999). Gap detection thresholds as a function of tonal duration for younger and older listeners. *J Acoust Soc Am*, *106*, 371–380.

Schwarz, D. W., & Taylor, P. (2005). Human auditory steady state responses to binaural and monaural beats. *Clin Neurophysiol*, *116*, 658–668.

Shamma, S. A., & Micheyl, C. (2010). Behind the scenes of auditory perception. *Curr Opin Neurobiol*, *20*, 361–366.

Shannon, R. V., Zeng, F. G., Kamath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science*, *270*, 303–304.

Siveke, I., Ewert, S. D., Grothe, B., et al. (2008). Psychophysical and physiological evidence for fast binaural processing. *J Neurosci*, *28*, 2043–2052.

Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear Hear*, *31*, 302–324.

Smith, J. C., Marsh, J. T., Brown, W. S. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol*, *39*, 465–472.

Snowling, M. J., & Göbel, S. M. (2011). Reading development and dyslexia. In Goswami, U. (Ed.), *Wiley-Blackwell Handbook of Childhood Cognitive Development* (2nd ed., pp. 524-548). Malden, MA: Wiley-Blackwell.

Snyder, J. S., Alain, C., Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *J Cogn Neurosci*, *18*, 1–13.

Snyder, J. S., Gregg, M. K., Weintraub, D. M., et al. (2012). Attention, awareness, and the perception of auditory scenes. *Front Psychol*, *3*, 15.

Snyder, R. L., & Schreiner, C. E. (1984). The auditory neurophonic: Basic properties. *Hear Res*, *15*, 261–280.

Starr, A., Picton, T. W., Sininger, Y., et al. (1996). Auditory neuropathy. *Brain*, *119*, 741–753.

Starr, A., McPherson, D., Patterson, J., et al. (1991). Absence of both auditory evoked potentials and auditory percepts dependent on timing cues. *Brain*, *114*, 1157–1180.

Starr, A., Zeng, F. G., Michalewski, H. J., et al. (2008). Perspectives on auditory neuropathy: Disorders of inner hair cell, auditory nerve, and their synapse. In A. I. Basbaum, A. Kaneko, G. M. Shepherd et al. (Series Eds), P. Dallos & D. Oertel (Volume Ed). *The Senses: A Comprehensive Reference, Vol 3, Audition* (pp. 397–412). San Diego, CA: Academic Press.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia*, *7*, 12–36.

Strong, G. K., Torgerson, C. J., Torgerson, D., et al. (2011). A systematic meta-analytic review of evidence for the effectiveness of the 'Fast ForWord' language intervention program. *J Child Psychol Psychiatry*, *52*, 224–235.

Sussman, E., Ritter, W., Vaughan, H. G., Jr. (1998). Predictability of stimulus deviance and the mismatch negativity. *Neuroreport*, *9*, 4167–4170.

Swaminathan, J., Krishnan, A., Gandour, J. T., et al. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Trans Biomed Eng*, *55*, 281–287.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang*, *9*, 182–198.

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nat Rev Neurosci*, *5*, 721–728.

Tremblay, K. L., Friesen, L., Martin, B. A., et al. (2003). Test-retest reliability of cortical evoked potentials using naturally produced speech sounds. *Ear Hear*, *24*, 225–232.

Tremblay, K., Kraus, N., McGee, T., et al. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear Hear*, *22*, 79–90.

Tremblay, K., & Moore, D. (2012). Current issues in auditory plasticity and auditory training. In K. Tremblay & R. Burkard (Eds). *Translational Perspectives in Auditory Neuroscience: Special Topics* (pp. 165–189). San Diego, CA: Plural Publishing.

Tremblay, K. L., Shahin, A. J., Picton, T., et al. (2009). Auditory training alters the physiological detection of stimulus-specific cues in humans. *Clin Neurophysiol*, *120*, 128–135.

Ungan, P., Ya cio lu, S., Ozmen, B. (1997). Interaural delay-dependent changes in the binaural difference potential in cat auditory brainstem response: Implications about the origin of the binaural interaction component. *Hear Res*, *106*, 66–82.

Vandermosten, M., Boets, B., Luts, H., et al. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proc Natl Acad Sci USA*, *107*, 10389–10394.

Viemeister, N. F., & Plack, C. (1993). Time analysis. In W. Yost, A. Popper, & R. Fay (Eds). *Human Psychophysics* (pp. 116–154). New York, NY: Springer-Verlag.

Wang, A. L., Mouraux, A., Liang, M., et al. (2008). The enhancement of the N1 wave elicited by sensory stimuli presented at very short inter-stimulus intervals is a general feature across sensory systems. *PLoS ONE*, *3*, e3929.

Werner, L. A., Folsom, R. C., Mancl, L. R., et al. (2001). Human auditory brainstem response to temporal gaps in noise. *J Speech Lang Hear Res*, *44*, 737–750.

Wible, B., Nicol, T., & Kraus, N. (2005). Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain*, 128, 417–423.

Won, J. H., Clinard, C. G., Kwon, S., et al. (2011). Relationship between behavioral and physiological spectral-ripple discrimination. *J Assoc Res Otolaryngol*, *12*, 375–393.

Wong, P. C., Skoe, E., Russo, N. M., et al. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci*, *10*, 420–422.

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

Yost, W. A. (1996). Pitch strength of iterated rippled noise. *J Acoust Soc Am*, *100*, 3329–3335.

Yost, W. A., Wightman, F. L., Green, D. M. (1971). Lateralization of filtered clicks. *J Acoust Soc Am*, *50*, 1526–1531.

Zeng, F. G., Kong, Y. Y., Michalewski, H. J., et al. (2005). Perceptual consequences of disrupted auditory nerve activity. *J Neurophysiol*, *93*, 3050–3063.