Auditory Attention Causes Gain Enhancement and Frequency Sharpening at Successive Stages of Cortical Processing—Evidence from Human Electroencephalography

Jessica de Boer and Katrin Krumbholz

Abstract

Previous findings have suggested that auditory attention causes not only enhancement in neural processing gain, but also sharpening in neural frequency tuning in human auditory cortex. The current study was aimed to reexamine these findings. Specifically, we aimed to investigate whether attentional gain enhancement and frequency sharpening emerge at the same or different processing levels and whether they represent independent or cooperative effects. For that, we examined the pattern of attentional modulation effects on early, sensory-driven cortical auditory-evoked potentials occurring at different latencies. Attention was manipulated using a dichotic listening task and was thus not selectively directed to specific frequency values. Possible attention-related changes in frequency tuning selectivity were measured with an adaptation paradigm. Our results show marked disparities in attention effects between the earlier N1 deflection and the subsequent P2 deflection, with the N1 showing a strong gain enhancement effect, but no sharpening, and the P2 showing clear evidence of sharpening, but no independent gain effect. They suggest that gain enhancement and frequency sharpening represent successive stages of a cooperative attentional modulation mechanism that increases the representational bandwidth of attended versus unattended sounds.

INTRODUCTION

There is manifold evidence that attention causes top-down modulation of sensory-driven or “exogenous” cortical responses (e.g., Fritz, Shamma, Elhilali, & Klein, 2003; Spitzer, Desimone, & Moran, 1988; Hillyard, Hink, Schwent, & Picton, 1973; reviewed in Fritz, Elhilali, David, & Shamma, 2007; Mangun & Hillyard, 1995), but the mechanisms underlying this modulation still remain unclear. Two alternative models have been proposed: The “gain enhancement” model assumes that attention increases neuronal responsiveness to the attended stimulus (McAdams & Maunsell, 1999; Hillyard, Vogel, & Luck, 1998), and the “sharpening” model assumes that attention increases neuronal tuning selectivity (Atiani, Elhilali, David, Fritz, & Shamma, 2009; Fritz et al., 2003; Spitzer et al., 1988). The current study aimed to test whether or how these models apply to the auditory domain. In particular, we wanted to test (i) whether exogenous auditory responses are really affected by attentional sharpening and, if so, (ii) how gain enhancement and sharpening relate within the context of the auditory processing hierarchy: Do they occur at the same or different processing levels, and do they operate cooperatively or independently of one another?

Numerous earlier studies have found noninvasively recorded auditory cortical responses to be larger when the evoking sound is attended, rather than unattended (electroencephalography (EEG)/magnetoencephalography: Fujiwara, Nagamine, Imai, Tanaka, & Shibasaki, 1998; Hillyard et al., 1998; Woldorff et al., 1993; Woldorff & Hillyard, 1991; Hillyard et al., 1973; fMRI: Jäncke, Mirazzade, & Shah, 1999), and have generally interpreted this finding within the context of a gain enhancement mechanism. More recently, however, it has been suggested that auditory attentional modulation also involves sharpening (Ahveninen et al., 2011; Kauramaki, Jääskeläinen, & Sams, 2007; Okamoto, Stracke, Wolters, Schmael, & Pantev, 2007). To demonstrate sharpening, the previous studies have used paradigms involving “notched noise” (NN) masking, a technique that has been used extensively in behavioral measurements of auditory frequency selectivity (e.g., Glasberg & Moore, 1990). NN masking requires the participant to attend to a fixed-frequency tone while trying to ignore a concurrently presented broadband noise with a spectral notch centered on the tone frequency. When the notch is narrow enough so that the tone response is partially obscured or “masked” by the noise response, the size of the unobscured portion of the tone response (over and above the noise response) should depend on the tuning selectivity of the tone-responsive neurons (Sams & Salmelin, 1994) and
should thus be sensitive to any sharpening in tuning selectivity induced by attention. Consistent with this expectation, the previous studies have found greater attentional enhancement of the tone response size when the notch was narrower than when it was wider (Kauramaki et al., 2007; Okamoto et al., 2007) or when the masking noise was omitted altogether (Ahveninen et al., 2011). Arguably, however, this finding could also be explained in terms of gain enhancement. This is because the tone was presented at a fixed intensity and would thus have been less audible when presented within a narrow notch. As a result, the unattended tone response size would have been smaller, and the attentional task would have been more difficult to perform. Earlier findings (Boudreau, Williford, & Maunsell, 2006; Alho, Woods, Algazi, & Näätänen, 1992; Schwent, Hillyard, & Galambos, 1976a, 1976b) suggest that both factors should have led to greater attentional gain enhancement, thus mimicking the effect of attentional sharpening.

To avoid these confounds, the current study manipulated attention and measured tuning selectivity independently using dichotic listening and adaptation, respectively.

Figure 1. Experimental design and neuron population model predictions. (A) Dichotic attention task: One ear received rapid sequences of sine tones varying pseudorandomly between four different frequencies (“Ipsi”), and the other received sequences of waxing noises (“Contra”). Participants were instructed to attend to either sequence alternately (~2 min cycle) and detect rare target sounds (shown in red; nontargets are shown in black) in the attended sequence. The tone targets were distinguished by an upward-sweeping frequency trajectory, and the noise targets were distinguished by a time-reversed (waning) amplitude envelope (see sound examples on the right; tone frequencies and stimulus durations not to scale). (B) Neuron population model predictions of attentional gain enhancement and sharpening (see Methods). The top and middle rows show the predicted effects of gain enhancement and sharpening separately, and the bottom row shows the effect of sharpening combined with a gain increase to counteract suppression in the aggregate response size. The left and middle columns show simulated population activity distributions to the adapting and adapted tones (open and filled outlines), plotted as a function of neuron characteristic frequency in quasilogarithmic (cochlear frequency) units. The right column shows the aggregate sizes of the adapted responses as a function of the adaptor-probe frequency separation. The unattended responses are shown in gray (replotted across rows), and the attended responses are shown in color.
Briely, Sumner, & Krumbholz, 2013), the degree of adaptation between two successive tones should depend on the degree of overlap between the neuron populations responsive to the tones, and this, in turn, should depend on the neurons’ frequency tuning selectivity. Figure 1B shows predictions of how the adapted tone response sizes might be affected by attentional gain enhancement and sharpening effects. The predictions are based on a simple neuron population model, with model neurons tuned for frequency and subject to activity-dependent adaptation or fatigue (see Methods for model details). Because of adaptation, the aggregate population response size to the current tone is predicted to increase with increasing frequency separation of the preceding tone, regardless of attention condition (Figure 1B, right). Under the assumption of a pure gain enhancement mechanism (with multiplicative gain; Figure 1B, top row), attention is predicted to increase the population response size equally across all frequency separations (if response size is expressed in logarithmic units), leaving the shape of the response size function unchanged. In contrast, a pure sharpening mechanism (Figure 1B, middle row) is predicted to increase the initial slope of the response size function (at small frequency separations) but, also, to cause an overall suppression in response size across all frequency separations. The suppression arises, because as the neurons’ tuning selectivity increases, fewer neurons are activated and thus the aggregate population response size decreases. To avoid suppression, the sharpening has to be combined with a gain enhancement such that the aggregate response size remains constant (Figure 1B, bottom row). As a result, the initial slope of the response size function is again predicted to steepen, but the response size now remains unchanged at zero and large frequency separations (when the responses to the successive tones overlap either completely or not at all; see Figure 1B, left and middle).

The previous studies that have used NN masking to investigate auditory attentional modulation mechanisms (Ahveninen et al., 2011; Kauramaki et al., 2007; Okamoto et al., 2007) have focused exclusively on the prominent N1 deflection of the CAEPs (Naätänen & Picton, 1987). Here, we also examined the preceding and following P1 and P2 deflections, which, like the N1, are exogenous and thus presumably represent earlier and later stages of sensory-driven auditory processing. Our results suggest that gain enhancement and sharpening represent cooperative components of a hierarchically distributed auditory attentional modulation mechanism, affecting different sensory-driven processing levels: The earliest observed attention effects (in the N1) appeared to be pure gain enhancement effects, whereas sharpening effects appeared to emerge only at later processing levels (in the P2). Our results suggest that gain enhancement and sharpening might work together to increase the representational bandwidth or “data rate” of attended over unattended auditory information.

METHODS
Participants
Twenty-three participants (seven men; mean age = 23.1, SD = 3.8 years) participated after having given written informed consent. All participants had hearing thresholds at or below 20 dB HL at all audiometric frequencies (250–8000 Hz) and had no history of audiological or neurological disease. The experimental procedures accorded with the Declaration of Helsinki (Version 6, 2008) and were approved by the Ethics Committee of the University of Nottingham School of Psychology, but were not formally preregistered online in accordance with the 2014 amendment to the declaration.

Stimuli and Procedure
During the EEG experiment, participants were comfortably seated in an electrically shielded, sound-attenuating booth (IAC Acoustics, Winchester, United Kingdom). The experiment consisted of four runs with short breaks in between. In three runs, referred to as “active runs,” participants were required to alternately attend to tone or noise sequences, presented to opposite ears, and detect infrequent targets within the attended ear. The to-be-attended ear was indicated by visual instruction and was switched every ~2 min. The ear of presentation of the tone and noise sequences was counterbalanced across participants. The active runs lasted about 12 min each. In the remaining run, referred to as “passive run,” the stimuli were presented passively while the participants watched a silent subtitled movie of their own choice to remain alert. The duration of the passive run was matched to the total duration for which participants attended to each ear over the three active runs (i.e., 3 × 6 min = 18 min). The active and passive runs were played consecutively, in counterbalanced order across participants.

The tones (“Ipsi” in Figure 1A) had a duration of 100 msec, including 20-msec cosine-squared onset and offset ramps, and were presented at a fixed stimulus onset interval (SOI) of 500 msec. A fixed SOI was used, because varying it would have varied the degree of adaptation between successive tones (Lanting et al., 2013) and thus confounded the tuning selectivity measurement. The tone frequencies were distributed equally between four different values, which were by 0, 75, 150, and 300 cents higher than 1000 Hz (1000, 1044, 1091, and 1189 Hz). The tone sequences were pseudorandom de Bruijn sequences consisting of 256 items each (lasting ~2 min). They were designed such that not only each frequency individually, but also each possible combination of two, three, or four consecutive frequencies occurred an exactly equal number of times (64, 16, 4, and 1, respectively; Brimijoin & O’Neill, 2010).

The noise stimuli (“Contra” in Figure 1A) were generated from equally exciting noise (with equal energy falling in each auditory filter; Glasberg & Moore, 2000),
which was box-car filtered between 2000 and 3000 Hz. They had a duration of 200 msec and were amplitude-modulated with a waxing amplitude envelope consisting of linear onset and offset ramps lasting 150 and 50 msec, respectively. The SOI of the noises was randomized between 666 and 966 msec (mean = 816 msec) to decorrelate the onset times of the tones and noises across the two ears.

The tone targets were distinguished from the non-target tones by a linearly rising frequency ramp (the non-target tones had a steady frequency; Figure 1A, right). They were presented randomly with a probability of 7.5%, with the constraint that every two successive target tones were separated by at least four non-target tones. The noise targets were time-reversed versions of the non-target noises (nontargets were waxing, and targets were waning noises; Figure 1A; idea taken from Cusack, Deeks, Aikman, & Carlyon, 2004). They were presented with a probability of 10% and separated by at least two non-target noises. On average, both the tone and noise targets occurred about 20 times within each ~2-min period (targets were presented within both the attended and unattended sequences).

All stimuli were generated digitally using MATLAB (The Mathworks, Natick, MA) and digital-to-analogue converted with a 24.414-kHz sampling rate and 24-bit amplitude resolution using TDT System 3 (Tucker Davis Technologies, Alachua, FL) consisting of an RP2.1 real-time processor and an HB7 headphone buffer. Both the tone and noise stimuli were presented at a sound pressure level of 65 dB using Sennheiser HD-280 Pro circumaural headphones (Sennheiser, Wedemark, Germany).

EEG Recordings

CAEPs were recorded with 33 Ag/AgCl ring electrodes (EASYCAP, Herrsching, Germany), placed according to the standard 10–20 layout, and a BrainAmp DC EEG amplifier (Brain Products, Gilching, Germany). Skin-to-electrode impedances were maintained below 5 kΩ throughout the recordings. The recording reference was the vertex (Cz) channel, and the ground was placed on the central forehead (AFz). The electrode signals were sampled at 500 Hz and bandpass-filtered online between 0.1 and 250 Hz using BrainVision Recorder (Brain Products). Only the responses to the nontarget tones were analyzed further.

EEG Data Analysis

The EEG data were first preprocessed using the EEGLAB toolbox (Delorme & Makeig, 2004), which runs under MATLAB. First, they were low-pass filtered at 35 Hz using a −48-dB/oct zero-phase IIR filter, and then they were re-referenced to average reference and segmented into 500-msec epochs ranging from 100 msec before to 400 msec after the onsets of the nontarget tones. Epochs containing unusually large amplitudes across electrodes (joint probability larger or equal to three standard deviations) were rejected automatically. The remaining epochs were submitted to an independent component analysis (extended infomax algorithm). Components representing eye blinks, lateral eye movements, and electrocardiac activity were removed by manual inspection of the components’ temporal traces and scalp topographies.

Activity during the baseline period of the tone responses (before the tone onset) was both highly non-stationary and also considerably larger for attended than unattended trials (Figure 2A), suggesting the presence of longer-lasting endogenous activity from preceding trials (Woldorff, 1993). To minimize the effect of this activity on the analysis of the discernible exogenous deflections (P1, N1, and P2; Figure 2A), we baseline-corrected each deflection separately, using a different baseline window (referred to as “deflection-specific” baseline correction). All windows were given a minimal duration of only 8 msec. The windows for the N1 and P2 were centered at the peaks of the respective preceding, opposite polarity deflections (P1 and N1, respectively), thus effectively creating a peak-to-peak difference. This would be expected to minimize any unipolar activity associated with endogenous attentional processing, such as the so-called “processing negativity” (Näätänen, 1990), which would affect opposite polarity deflections in opposite directions and thus cancel in the peak-to-peak difference. The window for the P1 was located at the tone onset (around 0 msec), close to the P1 deflection start. The baseline correction was performed separately for each participant and analyzed condition. The baseline-corrected deflections will be referred to as P10, N1P1, and P2N1 to indicate the differences in baseline window (see Figure 2B).

The P10, N1P1, and P2N1 peak amplitudes were measured both from the original sensor data and from source waveforms derived from source models fitted to each deflection peak. The sensor data were evaluated at the sensors that showed the largest unattended deflection peaks on average (Fz for the P10 and N1P1, and Cz for the P2N1) and referenced to the linked mastoids (average of TP9 and TP10). The source models were fitted to the unattended conditions only (when participants attended to the noise sequences in the opposite ear or watched a silent movie) to create a spatial filter for exogenous auditory cortical activity. They were implemented in the BESA software, version 5.3 (BESA, Gräfelfing, Germany), and each consisted of two hemispherically symmetric regional equivalent current dipoles (ECDs) (Scherg & Ebersole, 1993), with a four-shell ellipsoidal volume conductor as head model. First, the ECD locations were fitted to a 30-msec window centered at the relevant deflection peak in the grand-averaged response across all participants and unattended conditions. Then, the ECDs were realigned individually for each participant to maximize the peak source strength along their first dipole direction, and the resulting realigned first
dipole directions were used to extract source waveforms for each individual and condition. The source waveforms showed no significant hemispheric differences and were thus averaged across hemispheres.

The P10, N1P1, and P2N1 peak amplitudes were either averaged across all tone frequencies or evaluated separately for each absolute frequency separation, ΔF, between the current and preceding tones, which could take one of four values (0, 75, 150, or 300 cents). On average, the number of trials available for each absolute frequency separation and each participant was 391 (range: 347–414), 479 (range: 409–517), 481 (range: 428–507), and 241 (range: 208–258) when participants attended to the tone sequences, 397 (range: 374–419), 490 (range: 463–512), 491 (range: 465–508), and 245 (range: 231–260) when they attended to the noise sequences, and 380 (range: 330–406), 469 (range: 421–499), 467 (range: 412–506), and 236 (range: 212–248) when they watched a silent movie (passive run).

**Statistical Analyses**

Statistical analyses were conducted using R (R Core Team, 2013). Both the behavioral (hit/false alarm rates and RTs for target detection) and CAEP data (deflection peak amplitudes) were evaluated with linear mixed-effects models (nlme package; Pinheiro, Bates, DebRoy, Sarkar, & Team, 2017). The CAEP peak amplitudes were first converted to logarithmic units.

Homogeneity of variance was tested using Levene’s test (car package; Fox & Weisberg, 2011) and normality using quantile–quantile plots of the model residuals. Where variance homogeneity was violated (i.e., the residuals were significantly different across factor levels), each observation was weighted by the inverse of the variance for the respective factor level. This reduces the influence of noisier data points on the model fit. Normality was achieved by log transformation (applied to the false alarm rates and RTs). Any overly influential data points were identified using Cook’s distance and excluded.

In the models of the CAEP peak amplitudes, the linear frequency separation covariate (ΔF) was shifted downward by 150 cents (ΔF → ΔF − 150 cents) to reduce colinearity with the quadratic covariate (ΔF²). Next to the fixed effects, all models also contained by-subjects random intercepts and fixed-factor slopes. The fixed effects were fitted using maximum likelihood estimation and the random effects using restricted maximum likelihood estimation. Random effects were tested using log-likelihood ratio tests. Random effects that failed to produce a significant improvement in model fit were omitted. Fixed effects were evaluated using conditional F tests following the strategy described in Pinheiro and Bates (2000). Despite some missing data points, the number of data points was sufficiently similar across the various combinations of factor levels to allow Type III (marginal) tests to be evaluated for all included fixed effects. Significant fixed effects were post hoc tested using Tukey’s honestly significant difference (multcomp package; Hothorn, Bretz, & Westfall, 2008).

**Neuron Population Model of Attentional Modulation Effects**

Predicted effects of gain enhancement and sharpening were derived with a neuron population model, where...
each neuron was tuned to a different characteristic frequency, $f_c$, and subject to activity-dependent adaptation or fatigue, $A$. The shape of the frequency tuning was defined by a rounded-exponential (roex) function,

$$W(f) = \left(1 - p \frac{|f - f_c|}{f_c}\right) e^{-(|f - f_c|/p)^2} \quad (1)$$

where $f$ is the frequency of the tone stimulus, $|f - f_c|$ is its absolute separation from the neuron’s characteristic frequency ($f_c$), and $p$ is a parameter that determines the tuning sharpness. $p$ was set such that tuning curve widths corresponded to the equivalent rectangular bandwidths (ERBs) of the auditory frequency filters as determined by behavioral NN masking: $p = e^{rac{1}{\text{ERB}(f_c)}}$, where ERB ($f_c$) = 24.7(4.37$c + 1$) (Glasberg & Moore, 1990). The characteristic frequencies ($f_c$) varied from 50 to 16000 Hz and were distributed evenly on a cochlear frequency (ERB rate; Moore & Glasberg, 1983) scale.

Adaptation was modeled by multiplying the unadapted response to the current tone frequency, $f_0$ (given by $W(f_0)$; Equation 1) with a factor $1 - A$, where $A$ was proportional to the preceding tone frequency, $f_{-1}$ (given by $W(f_{-1})$). The degree of adaptation ($A$) was assumed to decay exponentially over time ($t$): $A(t) = A(t = 0) e^{-\tau t}$. The decay time constant, $\tau$, was set to 721.34 msec (compare Briley & Krumbholz, 2013; Roth et al., 1976), which meant that, between successive tone onsets, adaptation decayed by 50% (because $e^{-500/721.34} = 0.5$). The aggregate response size was derived by summing the adapted single-neuron responses across neurons.

Attentional gain enhancement was modeled by multiplying the single-neuron tuning functions $W$ (Equation 1) with a gain factor, $G > 1$. In the simulation shown in Figure 1B (upper row), $G$ was set to 2—doubling the attended compared with unattended response size. Attentional sharpening was modeled by dividing the tuning sharpness parameter, $p$, by a sharpening factor, $S < 1$. In the simulations shown in Figure 1B (middle and bottom rows), $S$ was set to 0.5—halving the ERBs of the attended compared with unattended tuning functions. If no gain is applied ($G = 1$), halving the ERBs halves the aggregate response sizes (middle row). To preserve the aggregate response size (bottom row), $G$ was concurrently raised to 2.

**RESULTS**

**Behavioral Results**

During the EEG recordings, participants either ignored the experimental sounds and watched a silent subtitled movie or alternately monitored the tone or noise sequences in the different ears for occasional target sounds (frequency-modulated tones and waning noises, respectively; Figure 1A). To match the difficulty in detecting the tone and noise targets, each participant first attended a short pilot session, where the target salience (determined by the frequency or amplitude modulation depth, respectively) was adjusted to yield a ~75% hit rate. Across participants, the adjusted frequency modulation depth of the tone targets ranged between 100 and 200 cents, and the amplitude modulation depth of the noise targets ranged between 50% and 100%.

During the experiment proper, the tone targets yielded an actual hit rate close to the adjusted rate (mean ± standard error: 76.0 ± 3.1%) and a false alarm rate of 10.0 ± 2.8%. In contrast, the actual hit rate for the noise targets was significantly higher (85.0 ± 2.8%; $F(1, 105) = 11.3$, $p = .0011$; here and onward, statistical tests are based on linear mixed-effects models [LMMs], with $F$ and $p$ values based on conditional $F$ tests; see Methods), and the false alarm rate was significantly lower (4.4 ± 1.3%; $F(1, 105) = 5.8$, $p = .0180$). At the same time, however, the noise targets also yielded a longer RT (613.6 ± 26.3 msec vs. 566.0 ± 18.0 msec for the tone targets; $F(1, 104) = 5.7$, $p = .0185$), suggesting that participants traded response speed for response accuracy. In the case of the tone sequences, the slope for such speed-accuracy trade-off was limited by the shorter SOI (500 msec vs. 816 ± 150 msec for the noise sequences; see Methods), which limited the RT. The presence of speed-accuracy trade-off is supported by the inverse efficiency score (IES), which combines response speed and accuracy measures into a single, overall measure of task performance (IES = RT/[1 − PE]), where RT is the reaction time and PE is the proportion of errors, i.e., false alarms and missed targets; Townsend & Ashby, 1978) and which was not significantly different between the tone and noise sequences (tones vs. noises: 814.4 ± 124.1 msec vs. 829.5 ± 72.2 msec; $F(1, 103) = 0.63$, $p = .5949$). The IES was also not significantly different across the three successive ~12-min measurement runs ("active runs"); main effect of run: $F(1, 103) = 0.63$, $p = .4274$; interaction between run and sequence type: $F(1, 103) = 0.28$, $p = .5949$).

### Average CAEPs

The average CAEPs to the nontarget tones (averaged across all frequency separations between successive tones; Figure 2A) exhibited three successive transient deflections P1, N1, and P2, which were clearly discernible and peaked at similar latencies (around 60, 105, and 150 msec) both when the tones were attended (Figure 2A, top) and when they were unattended (i.e., when participants attended to the noise sequences in the other ear or watched a silent movie; Figure 2A, bottom). Because of the relatively short SOI used (500 msec), the CAEPs failed to return to a steady baseline before the subsequent tone onset. As a result, the transient deflections were riding on a background of slowly varying nonstationary EEG activity from previous trials, which appeared to be particularly evident in the attended condition (Figure 2A, top). The nonstationarity of this background activity meant that it could not be eliminated by conventional baseline correction, and the use of a fixed
SOI (required to control the degree of adaptation between successive tones) meant that it could also not be eliminated by deconvolution-based methods (Lütkenhöner, 2010; Woldorff, 1993). To address this problem, we here opted to baseline-correct each deflection separately, using a baseline window that was both minimal in duration and located close to the respective deflection start (deflection-specific baselining; see Methods). The N1 and P2 were baseline-corrected to the respective preceding, opposite polarity peak—effectively creating a peak-to-peak difference. This would have minimized both the slowly varying previous-trial baseline, as well as any unipolar endogenous attention-related activity elicited within the current trial (such as the processing negativity; Näätänen, 1990). The baseline-corrected deflections are shown in Figure 2B (separately for each attention condition) and will be referred to as P10, N1P1, and P2N1. Figure 3A shows that they exhibited scalp voltage distributions typical of sources in supratemporal auditory cortex (characterized by a voltage inversion over the temporal bone; Scherg, Vajsar, & Picton, 1989; Vaughan & Ritter, 1970).

CAEPs measured at individual sensors may reflect a mixture of contributions from both exogenous and endogenous sources, but only the exogenous contributions represent the modulatory attention effects that we aim to investigate. Thus, to maximize these contributions, we analyzed the CAEPs not only in the original sensor space (using the sensors that showed the largest unattended peak amplitude for the respective analyzed deflection; see Methods and Figure 2B), but also in a source space representing exogenous sources. A different source model was used for each participant and analyzed deflection based on equivalent current dipoles fitted to the respective deflection peak in the individual unattended responses (where endogenous contributions should have been minimal; see Methods). Figure 3B shows that the best-fitting sources for all three unattended deflections (P10, N1P1, and P2N1) localized to the approximate auditory cortex region and that their average orientations were roughly perpendicular to the supratemporal plane. The goodness of fit ranged between 89.4% and 98.5% for the P1 (mean ± standard deviation: 96.4 ± 2.4), between 93.7% and 98.5% for the N1 (97.1 ± 1.4), and between 88.3% and 98.3% for the P2 (95.6 ± 2.2). The sources were used as spatial filters to extract source waveforms for each individual and condition (see Figure 3C for the grand-averaged source waveforms for each attention condition), and the source waveforms were averaged across hemispheres, because no significant hemisphere-specific condition effects were found.
Attention Effects on Average CAEPs

Comparison of the average CAEP waveforms between attention conditions (see Figures 2B and 3C for the sensor and source waveforms, respectively) suggests that the N1P1 and, to a lesser degree, also the P2N1 were enhanced in the attended compared with unattended, whereas the P10 seemed to be largely unaffected by attention. The waveforms also suggest that there was little difference between the two unattended conditions (i.e., when participants attended to the noise sequences or watched a silent movie, labeled “ignored” and “passive” in Figures 2B and 3C)—for any of the three deflections.

These results were confirmed by submitting the average deflection peak amplitudes (in logarithmic units; Figure 4A) to linear mixed-effects models (LMMs), with attention condition and deflection (if appropriate) as fixed factors. The models were calculated either for successive deflection pairs (P10/N1P1 and N1P1/P2N1) or for each deflection separately (henceforth referred to as “combined” or “separate LMMs”). Effects that were significant in the current but not the preceding deflection were interpreted as “emerging” at the level of the current deflection. Both for the sensor and for the source data, the combined LMM of the P10/N1P1 peak amplitudes revealed a significant overall (main) effect of attention condition (sensor: $F(2, 107) = 3.8, p = .0254$; source: $F(2, 105) = 6.3, p = .0026$) but also showed a significant deflection by attention condition interaction (sensor: $F(2, 107) = 15.5, p < .0001$; source: $F(2, 105) = 11.1, p < .0001$). The interaction arose, because the attention condition effect was significant only for the N1P1 (shown by the respective separate LMMs; sensor: $F(2, 42) = 33.0, p < .001$; source: $F(2, 42) = 28.7, p < .0001$) but nonsignificant for the P10 (sensor: $F(2, 43) = 0.8, p = .4717$; source: $F(2, 41) = 0.4, p = .6672$). This suggests that the attention condition effect first emerged at the level of the N1. In the combined LMM of the N1P1/P2N1 peak amplitudes, the main effect of attention condition was again significant for both the sensor and the source data (sensor: $F(2, 107) = 13.3, p < .001$; source: $F(2, 105) = 11.5, p < .0001$). In this case, the deflection by attention condition interaction was significant for the source data, $F(2, 105) = 5.7, p = .0045$, but nonsignificant for the sensor data, $F(2, 107) = 0.7, p = .5079$. Consistent with this, the separate LMM for the P2N1 showed a significant attention condition effect for the sensor data, $F(2, 42) = 4.1, p = .0233$, but not for the source data, $F(2, 41) = 2.12, p = .1274$. This suggests that the attention condition effect on the average P2N1 peak amplitudes was more labile than for the N1P1, and may reflect previous-trial activity. For the N1P1, the attention condition effect was due to larger
peak amplitudes in the attended compared with both unattended (ignored and passive) conditions. This was true for both the sensor (both $p \leq .0001$) and source data (both $p < .0001$). For the P2n1 sensor amplitudes, the difference between the attended and ignored conditions was significant ($p = .00971$), but the difference between the attended and the passive conditions was nonsignificant ($p = .22429$; see stars in Figure 4A). The ignored and passive conditions showed little or no differences between one another—for any deflection and in either the sensor or source data (all $ps > .4$).

Attention Effects on Frequency-specific Adaptation

To test whether the observed attention effects on the average deflection peak amplitudes were generated by gain enhancement or sharpening, we evaluated the peak amplitudes separately for the different frequency separations, $\Delta F$, from the preceding tone, which were expected to cause different degrees of adaptation (Figure 1B). In the statistical models (LMMs), frequency separation was included both as a linear ($\Delta F$) and quadratic ($\Delta F^2$) fixed covariate because, based on the neuron population model predictions (Figure 1B), the linear covariate alone was not expected to be able to capture the effect of sharpening. For gain enhancement, the model predicted a constant increase in the response size across all frequency separations from the preceding tone. Statistically, this should create a main effect of attention condition, with no interaction with either frequency separation covariate ($\Delta F$ or $\Delta F^2$). In contrast, the sharpening mechanism was predicted to cause the response size function to become steeper at small frequency separations, thus making the function more nonlinear. Statistically, this should give rise to a significant interaction between $\Delta F^2$ and attention condition. The average peak amplitudes had shown no significant differences between the ignored and passive conditions for any deflection (see Figure 4A), and the same was also true for the peak amplitudes as a function of frequency separation (Figure 4B, C). Therefore, the ignored and passive conditions were now merged to form a single “unattended” condition.

The N1p1 and P2n1 peak amplitudes increased with increasing frequency separation (Figure 4B, C), as predicted by the neuron population model (compare Figure 1B). The corresponding (separate) LMMs revealed that this increase was significant for both the sensor (main effect of $\Delta F$; N1p1: $F(1, 241) = 13.8, p = .003$; P2n1: $F(1, 243) = 15.2, p = .0001$) and source data (N1p1: $F(1, 241) = 5.3, p = .0221$; P2n1: $F(1, 243) = 17.9, p < .0001$). These results indicate that the N1 and P2 were affected by frequency-specific adaptation. In contrast, the peak amplitudes for the P1o showed little or no change with frequency separation, for either the sensor (main effects of $\Delta F$ and $\Delta F^2$; both $F(1, 235) \leq 0.5, p \geq .4788$) or source data (both $F(1, 236) \leq 0.8, p \geq .3751$), suggesting that the P1 was either not adapted or that adaptation in the P1 was nonspecific to frequency.

Figure 4 (B and C, middle) suggests that attention increased the N1p1 peak amplitudes about equally across all frequency separations. This finding was statistically confirmed by the nonsignificance of the interactions between attention condition and both $\Delta F$ and $\Delta F^2$ in the separate LMM for the N1p1, which applied to both the sensor (both $F(1, 241) \leq 1.3, p \geq .2558$) and source data (both $F(1, 241) \leq 2.5, p \geq .1164$), and is consistent with the neuron population model predictions for gain enhancement (compare Figure 1B, top right). In contrast, the attention effect on the P2n1 peak amplitudes depended strongly on frequency separation, with little or no increase at the zero and largest frequency separations (0 and 300 cents; 100 cents correspond to 1 semitone) but large increases at the intervening frequency separations (75 and 150 cents; Figure 4B and C, right). This pattern is consistent with the neuron population model predictions for sharpening combined with a commensurate gain enhancement to preserve the aggregate response size (compare Figure 1B, bottom right). Statistically, it was confirmed by the significance of the interaction between attention condition and $\Delta F$ in the separate LMM for the P2n1, which, again, applied to both the sensor ($F(1, 243) = 11.0, p = .001$) and source data ($F(1, 243) = 5.0, p = .0264$). The interaction between attention condition and $\Delta F$ was nonsignificant (sensor: $F(1, 243) = 0.4, p = .5292$; source: $F(1, 243) = 0.2, p = .6822$).

The difference in the pattern of frequency separation-dependent attention effects between the N1p1 and P2n1 was statistically confirmed by the three-way interaction between deflection, attention condition, and $\Delta F^2$ in the corresponding combined LMM (N1p1/P2n1). This interaction, which was significant in the sensor data ($F(1, 506) = 8.2, p = .0045$) and approached significance in the source data ($F(1, 506) = 3.2, p = .0764$) suggests that sharpening emerges only at the level of the P2. In contrast to the N1p1 and P2n1, the P1o peak amplitudes showed no significant attention effects, at any frequency separation, as confirmed by the lack of significant interactions between attention condition and both $\Delta F$ or $\Delta F^2$ in the separate LMMs for the P1o (sensor: both $F(1, 235) \leq 0.1, p \geq .7699$; source: both $F(1, 236) \leq 1.4, p \geq .2427$).

DISCUSSION

The current results suggest that the earliest effects of auditory attentional modulation are mediated by a pure gain enhancement mechanism and that sharpening emerges only at later processing stages. In the current results, the earliest measured deflection, the P1—presumed to be generated in primary auditory cortex (Yvert, Crouzeix, Bertrand, Seither-Preisler, & Pantev, 2001; Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; Mäkelä, Hämäläinen, Hari, & McEvoy, 1994)—was little or not affected by attention. The subsequent N1 showed a strong
attention-related enhancement in average peak amplitude but no differential effects on frequency-specific adaptation, suggesting that the N1 was affected by a pure gain enhancement mechanism. In contrast, the latest measured deflection, the P2, showed a lesser enhancement in average peak amplitude but a marked increase in the degree of adaptation specificity. Predictions from a neuron population model showed that the pattern of the effects in the P2 was consistent with a sharpening in neural tuning selectivity, combined with a commensurate gain enhancement so that the overall response size remained unchanged.

These results are consistent with previous studies that have also found large attentional enhancements in N1 peak amplitude (Neelon, Williams, & Garell, 2006a, 2006b; Hillyard et al., 1975) but contradict the conclusion of the previous NN masking studies (Ahveninen et al., 2011; Kauramaki et al., 2007; Okamoto et al., 2007) that attentional enhancement of the N1 is caused by neuronal sharpening. In the NN studies, attention was directed to a specific frequency value, and the audibility of the attended stimulus was allowed to vary across conditions. As explained above (Introduction), this would likely have led to variation in the amount of attentional gain enhancement, in a way that would have mimicked the expected effect of sharpening (Boudreau et al., 2006; Alho et al., 1992; Schwent et al., 1976a, 1976b). In the current study, attention was directed to one or other ear, and stimulus audibility was fixed across conditions. Our results thus suggest that attention can sharpen selectivity for a feature (here, frequency) even when attention is not selectively focused on a specific feature value. A similar conclusion was reached by Murray and Wojciulik (2004), who used an adaptation paradigm to demonstrate attentional sharpening for visual orientation. In both our and Murray and Wojciulik’s studies, the feature in which sharpening was observed (frequency and visual orientation, respectively) was task-relevant (in Murray and Wojciulik’s study, participants had to detect a change in image orientation; in our study, they had to detect a small frequency modulation). It is thus possible that task relevance is a prerequisite for sharpening to occur.

The absence of significant attention effects in the earliest, P1, deflection in the current study is consistent with several previous studies (Neelon et al., 2006a, 2006b; Hillyard et al., 1973) that have also found no significant P1 attention effects. Other studies, however, that have used shorter SOIs, did find significant attention effects in the P1 and even earlier deflections (Woldorff et al., 1993; Woldorff & Hillyard, 1991), suggesting that the first emergence of attention effects is graded with attentional load.

The current finding of a small but significant (in the sensor data) attentional enhancement in the average peak amplitude contrasts with some previous CAEP studies that have found either no significant change (Hillyard et al., 1973) or even a reduction (Hansen & Hillyard, 1980) in the P2 as a result of attention. The reduction has been attributed to a separate unipolar deflection, termed the “processing negativity” or “Nd,” thought to reflect endogenous attention-related processes (Näätänen, 1990). Because of its negative polarity, the Nd would be expected to add to any modulatory enhancement of the N1 but diminish any enhancement of the P2. In the current study, this effect would have been minimized by the deflection-specific baselining procedure used (see Methods). Significant attentional enhancement of the P2 has also been found in intracranial recordings from the auditory temporal region (Neelon et al., 2006a, 2006b), where any influence of the Nd may have also been minimal. The Nd can be demonstrated by calculating the difference wave between attended and unattended responses. In the current study, this was precluded by the experimental design: Difference waves can only be meaningfully calculated when the previous-trial baseline activity in the attended and unattended responses is either the same on average (e.g., Hillyard & Münte, 1984; Hansen & Hillyard, 1983) or can be effectively corrected for (e.g., Woldorff, 1993; Woldorff & Hillyard, 1991). In the current study, attended and unattended trials were temporally separated into different blocks, and so, attended trials were always preceded by attended trials, and unattended trials were always preceded by unattended trials. As a result, the attended responses exhibited a substantially larger previous-trial baseline, on average, than the unattended responses. Correcting for the baseline was also not possible, as this requires a sufficiently variable SOI (Lütkenhöner, 2010; Woldorff, 1993). In the current study, the SOI had to be fixed to control the degree of adaptation between successive trials.

The N1 and P2 have often been viewed as part of the same component process (the so-called “N1–P2 complex”). However, the marked differences in the pattern of their observed attention effects suggests that, rather than representing a unitary complex, the N1 and P2 represent different hierarchical levels of exogenous auditory processing that play distinct functional roles in conscious sound perception. This is supported by previous findings showing that the N1 and P2 differ not only in source structure (Godey, Schwartz, de Graaf, Chauvel, & Liégeois-Chauvel, 2001; Lütkenhöner & Steinsträter, 1998; Hari et al., 1987; Hari, Kaila, Katila, Tuomisto, & Varpula, 1982) but also in functional properties, such as dependence on prior stimulation, general arousal, aging, and auditory training (Herrmann, Henry, Johnsrude, & OBoleser, 2016; Tremblay, Ross, Inoue, McClannahan, & Collet, 2014; Ross, Jamal, & Tremblay, 2013; Ross & Tremblay, 2009; Crowley & Colrain, 2004; Roth et al., 1976).

The effect of attention on adaptation or “repetition suppression” has been investigated by several previous studies—particularly in the visual domain and using fMRI (see Henson & Mouchlianitis, 2007, for a review). The results from these studies, however, have been mixed, with some studies finding similar repetition suppression in both attended and unattended conditions (Vuilleumier, Schwartz, Duhoux, Dolan, & Driver, 2005; Bentley, Vuilleumier, Thiel,
Driver, & Dolan, 2003), but others finding repetition suppression to be either reduced (Murray & Wojciulik, 2004) or absent in unattended conditions (Henson & Mouchlianitis, 2007; Yi, Kelley, Marois, & Chun, 2006; Eger, Henson, Driver, & Dolan, 2004; Yi, Woodman, Widders, Marois, & Chun, 2004). The previous studies compared responses to repeated versus different stimuli but, unlike the current study, did not vary the degree of stimulus difference. The current results suggest that the amount of unattended repetition suppression should depend on the relation between the degree of stimulus difference and neuronal tuning selectivity: If we had compared repeated versus different tones with only a single frequency separation, we would have observed similar attended and unattended repetition suppression, if the frequency separation had been greater than 150 cents, but reduced or absent unattended repetition suppression if the frequency separation had been equal to or smaller than 150 cents (see Figure 4B, C).

Previous studies from the visual (Summerfield, Wyart, Johnen, & de Gardelle, 2011; Summerfield, Tritschuh, Monti, Mesulam, & Egner, 2008) and auditory (Todorovic, van Ede, Maris, & de Lange, 2011; Wacongne et al., 2011) domains have demonstrated that repetition suppression is not only determined by the local stimulus context (locally preceding stimuli) but is also modulated by prior expectation, such that the amount of repetition suppression is reduced when stimulus repetition is unexpected. This is contrary to the idea of bottom–up neuronal fatigue and has been taken to suggest that repetition suppression may instead reflect the action of a hierarchical predictive coding mechanism, which combines bottom–up stimulus representations with prior, top–down stimulus expectations (e.g., Friston, 2005; Knill & Pouget, 2004). Within this predictive coding framework, it has been hypothesized that attention may modulate the top–down stimulus expectations—increasing expectation for attended over unattended stimuli (Friston, 2009; Rao, 2005). Several recent studies have interpreted their findings within the context of this hypothesis (Hsu, Hämäläinen, & Waszak, 2014; Chennu et al., 2013; Kok, Rahnev, Jehee, Lau, & de Lange, 2012). The current study, however, suggests an alternative or at least complementary explanation. This is because all stimuli and all stimulus transitions (including higher-order transitions between nonconsecutive stimuli) were perfectly balanced (see Methods) and thus presumably equally expected—and attention was also distributed equally across all stimuli. This excludes an explanation in terms of top–down expectation and instead suggests that attention modulates bottom–up representational properties.

The P2 amplitude showed little or no attention-related change unless the frequency separation from the preceding tone was either intermediate. According to the neuron population model predictions, this suggest that the P2 was affected by a combination of sharpening and gain enhancement and that the amount of gain enhancement matched the degree of sharpening, such that the overall response size remained unchanged. This suggests that gain enhancement and sharpening are distinct but cooperative components of a hierarchically distributed attentional modulation mechanism, which adaptively adjusts the representational bandwidth of auditory cortical processing in accordance with attentional demand. Sharpening increases representational resolution, but without a commensurate enhancement in gain, this would lead to decrease in representational accuracy (because fewer channels would be activated). By combining and matching gain enhancement and sharpening effects, the auditory system can increase representational resolution while, at the same time, maintaining representational accuracy. And by cascading the gain enhancement and sharpening effects across different processing levels—presumably with different limitations on representational resources (Ahissar & Hochstein, 2004), the system retains the ability to quickly switch attention to new or currently unattended sounds.

**Acknowledgments**

This work was funded by the MRC (Intramural Program grants MC_U135097128 and MC_UU00010/2). We thank Sarah Jane Gibbs for her help in participant screening and data collection and Oliver Zobay for assistance with the statistical analysis.

Reprint requests should be sent to Jessica de Boer, MRC Institute of Hearing Research at the School of Medicine, University of Nottingham, Science Road, University Park, Nottingham, NG7 2RD, UK, or via e-mail: Jessica.deBoer@nottingham.ac.uk.

**REFERENCES**

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8, 457–464.  
Ahveninen, J., Hämäläinen, M., Jääskeläinen, I. P., Ahlfors, S. P., Huang, S., Lin, F. H., et al. (2011). Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. *Proceedings of the National Academy of Sciences, U.S.A.*, 108, 4182–4187.  
Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology*, 82, 356–368.  
Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., & Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, 61, 467–480.  
Bentley, P., Vuilleumier, P., Thiel, C. M., Driver, J., & Dolan, R. J. (2003). Effects of attention and emotion on repetition priming and their modulation by cholinergic enhancement. *Journal of Neurophysiology*, 90, 1171–1181.  
Boudreau, C. E., Williford, T. H., & Maunsell, J. H. (2006). Effects of task difficulty and target likelihood in area V4 of macaque monkeys. *Journal of Neurophysiology*, 96, 2577–2587.  
Briley, P. M., Breakey, C., & Krumbholz, K. (2013). Evidence for pitch chroma mapping in human auditory cortex. *Cerebral Cortex*, 23, 2601–2610.
Briley, P. M., & Krumbholz, K. (2013). The specificity of stimulus-specific adaptation in human auditory cortex increases with repeated exposure to the adapting stimulus. *Journal of Neurophysiology, 110*, 2679–2688.

Brimijoin, W. O., & O’Neill, W. E. (2010). Patterned tone sequences reveal non-linear interactions in auditory spectrotimetral receptive fields in the inferior colliculus. *Hearing Research, 267*, 96–110.

Chenu, S., Noreika, V., Guerguev, D., Blennmann, A., Kochen, S., Ibáñez, A., et al. (2013). Expectation and attention in hierarchical auditory prediction. *Journal of Neuroscience, 33*, 11194–11205.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology, 115*, 732–744.

Cusack, R., Deeks, J., Aitken, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 643–656.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 15*, 9–21.

Edmonds, B. A., & Krumbholz, K. (2014). Are interaural time and level differences represented by independent or integrated codes in the human auditory cortex? *Journal of the Association for Research in Otolaryngology, 15*, 103–114.

Eger, E., Henson, R. N., Driver, J., & Dolan, R. J. (2004). BOLD repetition decreases in object-responsive ventral visual areas depend on spatial attention. *Journal of Neurophysiology, 92*, 1241–1247.

Fox, J., & Weisberg, S. (2011). An R companion to applied regression (2nd ed.). Los Angeles: Sage.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 360, 815–836.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences, 13*, 293–301.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—Focusing the searchlight on sound. *Current Opinion in Neurobiology, 17*, 457–455.

Fritz, J. B., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotimetral receptive fields in primary auditory cortex. *Nature Neuroscience, 6*, 1216–1223.

Fujiwara, N., Nagamine, T., Imai, M., Tanaka, T., & Shibasaki, H. (1998). Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer. *Brain Research. Cognitive Brain Research, 7*, 99–109.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*, 103–138.

Glasberg, B. R., & Moore, B. C. J. (2000). Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. *Journal of the Acoustical Society of America, 108*, 2318–2328.

Godøy, B., Schwartz, D., de Graaf, J. B., Chauvel, P., & Liégeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clinical Neurophysiology, 112*, 1850–1859.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14–23.

Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology, 49*, 277–290.

Hansen, J. C., & Hillyard, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 1–19.

Hari, R., Kaila, K., Katila, T., Tuomisto, T., & Varpula, T. (1982). Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: Implications for their neural generation. *Electroencephalography and Clinical Neurophysiology, 54*, 561–569.

Hari, R., Pelizzzone, M., Mäkelä, J. P., Hallström, J., Leinonen, L., & Lounasmaa, O. V. (1987). Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology, 26*, 31–43.

Henson, R. N., & Mouchlianitis, E. (2007). Effect of spatial attention on stimulus-specific haemodynamic repetition effects. *Neuroimage, 35*, 1317–1329.

Herrmann, B., Henry, M. J., Johnsrude, I. S., & Ohtsuki, W. (2016). Altered temporal dynamics of neural adaptation in the aging human auditory cortex. *Neurobiology of Aging, 45*, 10–22.

Hewson-Stoate, N., Schönwiesner, M., & Krumbholz, K. (2006). Vowel processing evokes a large sustained response anterior to primary auditory cortex. *The European Journal of Neuroscience, 24*, 2661–2671.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science, 182*, 177–180.

Hillyard, S. A., & Münte, T. F. (1984). Selective attention to color and location: An analysis with event-related brain potentials. *Attention, Perception, & Psychophysics, 36*, 185–198.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 353*, 1257–1270.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*, 346–365.

Hsu, Y. F., Hämaläinen, J. A., & Waszak, F. (2014). Repetition suppression comprises both attention-independent and attention-dependent processes. *Neuroimage, 98*, 168–175.

Jäncke, L., Mirrazazade, S., & Shah, N. J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: A functional magnetic resonance imaging study in human subjects. *Neuroscience Letters, 266*, 125–128.

Kauramaki, J., Jääskeläinen, I. P., & Sams, M. (2007). Selective attention increases both gain and feature selectivity of the human auditory cortex. *PLoS One, 2*, e909.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27*, 712–719.

Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in sensory processing. *Cerebral Cortex, 22*, 2197–2206.

Lanting, C. P., Briley, P. M., Sumner, C. J., & Krumbholz, K. (2015). Mechanisms of adaptation in human auditory cortex. *Journal of Neurophysiology, 110*, 973–983.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology, 92*, 204–214.
Lütkenhöner, B. (2010). Baseline correction of overlapping event-related responses using a linear deconvolution technique. NeuroImage, 52, 86–96.

Lütkenhöner, B., & Steinsträtter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiology & Neuro-Otology, 3, 191–213.

Magezi, D. A., & Krumholz, K. (2010). Evidence for opponent-channel coding of interaural time differences in human auditory cortex. Journal of Neurophysiology, 104, 1997–2007.

Mäkelä, J. P., Hämläinen, M., Hari, R., & McEvoy, L. (1994). Whole-head mapping of middle-latency auditory evoked magnetic fields. Electroencephalography and Clinical Neurophysiology, 92, 414–421.

Mangun, G. R., & Hillyard, S. A. (1995). Mechanisms and models of selective attention. In M. D. Rugg & M. G. H. Coles (Eds.), Electrophyiology of mind: Event-related brain potentials and cognition (pp. 40–85). New York: Oxford University Press.

McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. Journal of Neuroscience, 19, 431–441.

Moore, B. C. J., & Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. Journal of the Acoustical Society of America, 74, 750–753.

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. Nature Neuroscience, 7, 70–74.

Näätänen, R. (1990). The role of attention in auditory information-processing as revealed by event-related potentials and other brain measures of cognitive function. Behavioral and Brain Sciences, 13, 201–232.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology, 24, 375–425.

Neelon, M. F., Williams, J., & Garell, P. C. (2006a). The effects of attentional load on auditory ERPs recorded from human cortex. Brain Research, 1118, 94–105.

Neelon, M. F., Williams, J., & Garell, P. C. (2006b). The effects of auditory attention measured from human electrocorticograms. Clinical Neurophysiology, 117, 504–521.

Okamoto, H., Stracke, H., Wolters, C. H., Schmael, F., & Pantev, C. (2007). Attention improves population-level frequency tuning in human auditory cortex. Journal of Neuroscience, 27, 10983–10990.

Pinheiro, J., & Bates, D. (2000). Mixed-effects models in S and S-PLUS. New York: Springer.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2017). nlme: Linear and nonlinear mixed effects models. Retrieved from https://CRAN.R-project.org/package=nlme.

R Core Team. (2013). A language and environment for statistical computing. Retrieved from www.R-project.org.

Rao, R. P. (2005). Bayesian inference and attentional modulation in the visual cortex. NeuroReport, 16, 1843–1848.

Ross, B., Jamiäl, S., & Tremblay, K. L. (2015). Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. BMC Neuroscience, 14, 151.

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

Roth, W. T., Krainz, P. L., Ford, J. M., Tinklenberg, J. R., Rothbart, R. M., & Kopell, B. S. (1976). Parameters of temporal recovery of the human auditory evoked potential. Electroencephalography and Clinical Neurophysiology, 40, 625–632.

Salmén, N. H., May, P. J., Alku, P., & Tiitinen, H. (2009). A population rate code of auditory space in the human cortex. PLoS One, 4, e7600.

Sams, M., & Salmelin, R. (1994). Evidence of sharp frequency tuning in the human auditory cortex. Hearing Research, 75, 67–74.

Scherg, M., & Ebersole, J. S. (1993). Models of brain sources. Brain Topography, 5, 419–423.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. Journal of Cognitive Neuroscience, 1, 336–355.

Schwent, V. L., Hillyard, S. A., & Galambos, R. (1976a). Selective attention and the auditory vertex potential. Effects of signal intensity and masking noise. Electroencephalography and Clinical Neurophysiology, 40, 615–622.

Schwent, V. L., Hillyard, S. A., & Galambos, R. (1976b). Selective attention and the auditory vertex potential. I. Effects of stimulus delivery rate. Electroencephalography and Clinical Neurophysiology, 40, 604–614.

Snow, M., Coen-Cagli, R., & Schwartz, O. (2017). Adaptation in the visual cortex: A case for probing neuronal populations with natural stimuli. F1000Research, 6, 1246.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. Science, 240, 338–340.

Summerfield, C., Trüttschuh, E. H., Monot, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. Nature Neuroscience, 11, 1004–1006.

Summerfield, C., Wyart, V., Johnen, V. M., & de Gardelle, V. (2011). Human scalp electroencephalography reveals that repetition suppression varies with expectation. Frontiers in Human Neuroscience, 5, 67.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: An MEG study. Journal of Neuroscience, 31, 9118–9123.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), Cognitive theory (Vol. 3, pp. 199–239). Hillsdale, NJ: Erlbaum.

Tremblay, K. L., Ross, B., Inoue, K., McClannahan, K., & Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? Frontiers in Systems Neuroscience, 8, 28.

Vaughan, H. G., Jr., & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. Electroencephalography and Clinical Neurophysiology, 28, 360–367.

Vuilleumier, P., Schwartz, Z., Duhoux, S., Dolan, R. J., & Driver, J. (2005). Selective attention modulates neural substrates of repetition priming and “implicit” visual memory: Suppressions and enhancements revealed by fMRI. Journal of Cognitive Neuroscience, 17, 1245–1260.

Wacongne, C., Lahy, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. Proceedings of the National Academy of Sciences, U.S.A., 108, 20754–20759.

Webster, M. A. (2015). Visual adaptation. Annual Review of Vision Science, 1, 547–567.

Woldorf, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. Psychophysiology, 30, 98–119.
Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences, U.S.A.*, 90, 8722–8726.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, 79, 170–191.

Yi, D. J., Kelley, T. A., Marois, R., & Chun, M. M. (2006). Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Research*, 1080, 53–62.

Yi, D. J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, 7, 992–996.

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