



Unraveling the effects of selective auditory attention in ERPs: From the brainstem to the cortex

Daniel J. Strauss^{a,ib,*}, Farah I. Corona-Strauss^a, Adrian Mai^a, Steven A. Hillyard^{b,c}

^a Systems Neuroscience and Neurotechnology Unit, Faculty of Medicine, Saarland University & htw saar, Homburg/Saar, Germany

^b Department of Neurosciences, University of California San Diego, La Jolla, CA, USA

^c Leibniz-Institute for Neurobiology, Magdeburg, Germany

ARTICLE INFO

Dataset link: neuroimage-25-121295.snn-unit.de

Keywords:

Auditory attention
Brainstem
Event-related potentials
Neural synchrony
N1 effect

ABSTRACT

A little over fifty years ago, it was reported that selectively attending to one of two dichotically presented tone sequences enhances the major N1 component of the cortical event-related potential (ERP) to the attended tones. The present study revisited this classic experiment but replaced the tones in one ear with frequency-modulated “chirps” that were designed to activate the entire cochlea simultaneously and thereby elicit robust ERPs in the auditory brainstem pathways. Participants attended selectively to the sounds in one ear at a time with the task of reporting occasional targets of lower intensity. When chirps were attended, they elicited enhanced ERPs at multiple levels of the auditory pathways (0–250 ms), including a brainstem response at the level of the inferior colliculus. These results help to resolve a long-standing question of whether selective attention exerts top-down control over the initial transmission of competing auditory inputs in the brainstem pathways.

1. Background

The neural bases of auditory selective attention in humans have been investigated extensively by non-invasively recording auditory event-related potentials (ERPs), which can track stimulus-evoked neural activity all the way from the auditory nerve, through the brainstem relays, and ultimately to multiple levels of the auditory cortex (Picton, 2010). A critical design feature of such studies is the presentation of attended and ignored stimuli in unpredictable order to rule out possible confounding effects of non-specific processes on the neural response such as global arousal or alertness (Näätänen, 1967). The first electrophysiological study of auditory attention in humans to incorporate this paradigm was carried out by Hillyard and colleagues (Hillyard et al., 1973), who recorded ERPs to tone bursts presented to the right and left ears in random order. The major finding was that the prominent N1 component of the auditory ERP with a peak latency of 60–100 ms after stimulus onset was enhanced in response to attended-ear tones. This “N1 effect” was interpreted as an early selection of attended channel inputs for further processing in the manner of a sensory gain control mechanism. The neural generators of the enhanced N1 component and a longer-lasting “processing negativity” (Näätänen et al., 1978; Hansen and Hillyard, 1980; Giard et al., 2000) were localized to the auditory cortex (Scherg and Von Cramon, 1985; Näätänen and Picton, 1987; Woldorff et al., 1993). An earlier cortical component in the mid-latency range (20–50 ms) was also found to be modulated by attention prior the

N1 effect (Woldorff and Hillyard, 1991; Woldorff et al., 1993), lending support to the proposal that attention acts as a gain control over early evoked neural activity in auditory cortex.

What has remained unresolved in the 50-some years since the N1 effect was reported is whether selective attention can modulate auditory input at subcortical levels, which would indicate a mechanism of very early selection. The wealth of descending pathways from cortex to different levels of the auditory brainstem pathways could conceivably impose selective processing of attended stimuli in the brainstem (Winer et al., 1998; Suga et al., 2002; Blackwell et al., 2020), and experiments in animals have demonstrated such effects, see, e.g., Oatman and Anderson (1977). Numerous studies have investigated the possibility of attention effects at the brainstem level in humans through scalp recordings of the auditory brainstem evoked response (ABR, waves I–VI) elicited within the first 10 ms after a brief sound (click or tone) presentation. The broad consensus of these studies has been that the ABR is invariant to manipulations of attention (Picton et al., 1971, 1974; Woods and Hillyard, 1978; Hackley et al., 1987, 1990; Collet and Duclaux, 1986; Connolly et al., 1989; Gregory et al., 1989; Hirschhorn and Michie, 1990; Woldorff and Hillyard, 1991).

More recently, studies of the auditory frequency following response (FFR) to speech stimuli have revitalized the controversial question of whether attention can affect auditory transmission in the brainstem. The FFR is a near-sinusoidal oscillatory potential in the range of

* Corresponding author.

E-mail address: daniel.strauss@uni-saarland.de (D.J. Strauss).

<https://doi.org/10.1016/j.neuroimage.2025.121295>

Received 22 February 2025; Received in revised form 20 March 2025; Accepted 23 May 2025

Available online 7 June 2025

1053-8119/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

100–300 Hz that originates from the auditory midbrain pathways and is phase-locked to the voicing frequency of the speech input. Considerable evidence has accumulated that the scalp-recorded FFR to speech sounds shows increased amplitude and/or phase-locking when those sounds are attended (Galbraith et al., 1998; Lehmann and Schönwiesner, 2014; Forte et al., 2017; Etard et al., 2019).

It remains uncertain whether the attention-related modulation of speech-evoked neural activity in the brainstem is specific to the oscillatory FFR elicited by speech stimuli, or if it extends to a broader range of stimuli (Galbraith and Arroyo, 1993; Price and Bidelman, 2021). Moreover, the FFR has cortical contributions (Coffey et al., 2016), which are themselves influenced by attention (Schüller et al., 2023), complicating efforts to functionally unravel attentional modulations from the brainstem to the cortex. Importantly, these studies measured FFR amplitudes using either correlational (Forte et al., 2017) or frequency-domain (Lehmann and Schönwiesner, 2014) analyses, which are not well-suited for determining the precise time at which attention affects the auditory signal relative to stimulus onset. In particular, it is unclear whether the observed attentional enhancement of the FFR is initiated at stimulus onset or whether it requires time to build up, perhaps involving feedback from higher areas (Price and Bidelman, 2021). Accordingly, a key question left unresolved by the FFR studies (and previous ABR studies) is whether attention, under certain circumstances, can influence the initial feed-forward activity in the brainstem auditory pathways elicited by brief, non-sinusoidal sounds.

In the present study, we revisit this question by recording ABRs to tonal “chirp” stimuli in a dichotic listening design that paralleled that of the original N1 effect study by Hillyard and colleagues (Hillyard et al., 1973) to get full-range potentials (Kohl et al., 2019), mapping the hearing path from the brainstem to the cortex. Specifically, the chirp is a brief tone burst that is frequency-modulated from low-to-high in such a way that it activates hair cells along the entire length of the cochlea in synchrony, thereby eliciting pronounced ABRs at medium to low stimulation intensities (Dau et al., 2000; Corona-Strauss et al., 2009; Fobel and Dau, 2004; Elberling et al., 2010) (see Fig. S1 in Supplementary Information; SI). Participants were presented with randomized sequences of broadband chirps to one ear and 800 Hz tone bursts to the other ear with instructions to attend to one ear at a time and report occasional targets of lower intensity in the attended ear. Electroencephalographic (EEG) activity was recorded with vertex-mastoid configurations (“vertex” recordings) as used in the original study (Hillyard et al., 1973) as well as with a high-density (128-channel) electrode array (“multi-channel” recordings). Using different time-domain and wavelet-based time-frequency approaches, single-trial as well as averaged ERPs were analyzed with the aim of examining attentional modulations of auditory input at brainstem (ABR; 0–10 ms), mid-latency (auditory middle latency response, AMLR; 10–50 ms), and long-latency (auditory late response, ALR; 50–250 ms) levels. To preview the results, selective attention significantly modulated the neural activity evoked by transient auditory stimuli at multiple levels of the auditory pathways, beginning with wave V of the ABR.

2. Methods

2.1. Participants

Thirty-one normal hearing volunteers from the university environment participated in the study (26.8 ± 3.7 y, 23 m/8f). The inclusion criteria were a mean hearing level (HL) of <15 dB SPL across frequencies as measured via pure-tone audiograms (PTA) and no history of neurological diseases. Before collecting PTAs, participants were briefed on the experimental paradigm that was designed in line with the Declaration of Helsinki and approved by the local ethics committee (Ärztchamber des Saarlandes, Saarland Medical Council; Identification number: 305/20).

2.2. Stimuli and experimental paradigm

The auditory selective attention paradigm followed the original used by Hillyard et al. (1973) as closely as possible while incorporating some important modifications. In particular, stimuli were presented to the left and right ears in random order according to a single stimulation schedule. Stimuli in one ear consisted of 800 Hz tone bursts, while the other ear received a sequence of broadband (100–9800 Hz) chirps. Inter-stimulus intervals were randomized between 250–400 ms, and each successive stimulus event was assigned to the left or right ear at random. Participants were cued to attend to sounds in one ear at a time, and their task was to press a button each time an occasional target stimulus of lower intensity occurred.

Tone bursts lasted 50 ms including rise and fall times of 12 ms. The competing chirp stimuli were chosen because of their spectrotemporal characteristics, which produce a synchronized excitation along a larger portion of the cochlea, resulting in more pronounced ABRs at low to medium stimulation levels (Fobel and Dau, 2004). Specifically, chirps were designed according to the A-Chirp in Fobel and Dau (2004) and Corona-Strauss et al. (2009) for an intensity level of 40 dB SPL. Stimuli were created with a sampling rate of 44100 Hz, and their generation as well as all data acquisition and processing pipelines were implemented using software for scientific computing (Matlab and Simulink, The Mathworks, USA). The acoustic waveforms for both stimuli are shown in Fig. S1 in the SI along with additional information on the stimuli.

The total experiment consisted of 20 trials, each lasting 195.3 ± 1.2 s and presenting about 300.0 ± 25.4 stimuli to each ear. Target stimuli of 15 dB SPL lower intensity were randomly interposed every 3–20 stimuli with a 9.4% probability of occurrence. Experimental trials were constructed from ten different randomized scenarios, half with chirps in the left ear and half with chirps in the right ear. Each scenario was presented twice, under attend-left and attend-right conditions. Thus, overall there were five trials for each combination of stimulus type (chirps or tones), condition (attended or ignored), and location (left or right ear). The order of presentation of these trials was pseudo-randomized across participants.

Stimuli were presented via circumaural headphones (HDA300, Sennheiser, Germany) and playback was coordinated by a cross-platform digital audio workstation (StudioOne, PreSonus, USA) coupled to an audio interface (Scarlett 18i20, Focusrite, United Kingdom). The hearing thresholds for the chirps and tones were determined for each participant, and the intensity of the standard stimuli of both types was set to 40 dB sensation level. Finally, a fine-tuning process was carried out to balance the loudness perception across ears and stimulus types as described in the SI.

To familiarize participants with the stimuli and the behavioral task, a short training phase (approx. 3 min) was administered before the actual experiment. During the recording sessions, they sat in a comfortable armchair and were instructed to focus on a fixation ball at sight level at 2 m distance and to move as little as possible. Rest breaks were given between runs as necessary.

2.3. Electrophysiological recording

EEG recordings were obtained from two different sensor configurations. Similarly to the original straightforward recording scheme (Hillyard et al., 1973), five passive silver/silver-chloride (Ag/AgCl) electrodes were placed to record the EEG at the upper and lower mastoid behind each ear as well as at the vertex position. To complement this data with information of high spatial resolution, an additional EEG cap with 128 active sintered Ag/AgCl electrodes arranged in the 10–5 layout (g.SCARABEO, gtec, Austria) was placed on top of the passive electrodes. Furthermore, to ensure that none of the observed effects originated from activations of the postauricular muscle (PAM) due to selective attention (Strauss et al., 2020), a pair of passive electrodes was placed on each PAM to provide bipolar monitoring of

their electromyographic activity. All electrophysiological signals were recorded against a passive ground electrode on the upper forehead and sampled at 9600 Hz using a biosignal amplifier (g.HIAMP, gtec, Austria). Throughout the experiment, impedances for passive and active electrodes were controlled to be lower than 10 k Ω and 50 k Ω , respectively.

2.4. EEG preprocessing and segmentation

EEG preprocessing was conducted in two stages, with an initial stage for common processing of vertex and multi-channel recordings, and a second stage for artifact correction. The experimental paradigm was designed to reliably evoke auditory full-range ERPs including the ABR, the AMLR, and the ALR. Since ABRs are preferably investigated using a mastoid reference ipsilateral to stimulation (Picton, 2010), all recordings were referenced to the mastoid ipsilateral to attended as well as to ignored stimuli which resulted in two data sets per experimental trial. Data were then decimated to 4800 Hz, zero-phase comb-filtered at 50 Hz and its harmonics up to 2400 Hz, zero-phase bandpass-filtered at 1–1500 Hz, and corrected for DC-offsets. Multi-channel data from all experimental trials were then concatenated (including both ipsilaterally referenced data sets per trial) and the mean Pearson correlation between each electrode and its five nearest neighbors was calculated. Channels presenting correlations below two standard deviations from the mean across all channels were interpolated using EEGLAB's (DeLorme and Makeig, 2004) spherical interpolation algorithm (6.52 ± 1.90 channels per participant). Finally, data were split into the individual data sets and centered around 0 μ V.

ERPs for vertex recordings were extracted over epochs of -1000 ms to 1000 ms relative to onsets of standard stimuli; target ERPs were not analyzed to avoid inclusion of target detection-related components such as the P3 (Squires et al., 1975). Since the tone and chirp stimuli differed considerably in their physical properties, ERP onsets within the extracted segments were adjusted relative to stimulus-specific reference time points; specifically, time-zero for the ERP segments was chosen as the rise time offset at 12 ms for tones and stimulus offset at 14.9 ms for chirps (see SI section on Chirp-evoked ABRs for more details about the zero time reference). After establishing this reference point, latency values of ABR components are comparable to those elicited by click stimuli (Dau et al., 2000; Elberling and Don, 2008; Elberling et al., 2010). To remove signal offsets while considering the latencies of the earliest ERP components, single-trials were baseline-corrected by subtracting the mean potential over the interval -2 ms to 2 ms relative to the adjusted ERP onsets. Finally, for each combination of stimulus type (chirps or tones), condition (attended or ignored), and location (left or right ear), the first 1000 trials that did not exceed an absolute amplitude threshold of 100 μ V were selected and pooled across left and right ear locations for further analysis. The corresponding ERPs were extracted from the multi-channel recordings and baseline-corrected as reported above. Since the grand-average waveforms of vertex recordings revealed clear components in the ABR, AMLR, and ALR latency ranges for the chirps but not for the tones (see Fig. 1), all subsequent analyses were based on chirp-evoked responses.

2.5. ERP wavelet and spectrotemporal consistency analysis

Single-trials from vertex recordings were mapped to time–frequency representations by means of continuous wavelet transforms (CWTs) using analytic Morse wavelets (Lilly and Olhede, 2012). Specifically, data were analyzed at logarithmically spaced wavelet peak frequencies over 4–512 Hz with 32 scales per octave. Because of this broad spectral range, the number of wavelet cycles within the central time-domain power window was linearly increased over in $[0.75, 2]$ with 0.18 cycles increase per octave to simultaneously provide satisfactory temporal resolution at low frequencies and spectral resolution at high frequencies, e.g., see Morales and Bowers (2022). This was achieved

by fixing the wavelet family parameter at $\gamma = 3$ and varying the order β over 1.85–13.16. The complex CWT coefficients were then used to extract power- and phase-based time–frequency measures. Single-trial wavelet power was computed by squaring the absolute values of the coefficients and averaging the resulting power matrices across single-trials. The final power matrices were then baseline-normalized independently for each wavelet scale by dividing the power at each time point by the mean power over the interval -500 ms to -125 ms and log-transformed ($10\log_{10}$) to decibels. In addition, inter-trial phase coherence (ITPC) was calculated by extracting the instantaneous phase angle at each time–frequency point and calculating the mean resultant vector length (Rao Jammalamadaka and Sengupta, 2001) across single-trials in a point-wise manner.

2.6. ERP waveform consistency analysis

The ITPC analysis revealed several time–frequency clusters matching the spectrotemporal scales of the ABR, AMLR, and ALR within which selective attention significantly increased phase consistency across single-trials (see Fig. 3, top). Based on this observation, a subsequent analysis investigated whether attention also enhanced waveform consistency on a single-trial basis in the time-domain. In this analysis, narrowband bandpass filters were derived from the significant ($p < 0.01$) ITPC frequency ranges and applied to the averaged vertex and multi-channel ERPs as well as their respective single-trials. The (zero-phase) filtered signals were then trimmed to the characteristic latency ranges to extract the ABR (0–10 ms, 105.3–185.0 Hz), AMLR (10–50 ms, 44.3–71.3 Hz), and ALR (50–250 ms, 4.0–15.7 Hz). Finally, the waveform consistency was calculated as the mean Pearson correlation between the waveform of the averaged ERP and each of its single-trials for each of the three ERP latency ranges.

2.7. ABR enhancement and peak analysis

ERPs were subjected to four additional types of time-domain analyses to extract different ABR representations with enhanced signal-to-noise ratios. The first approach applied a conventional broad ABR (zero-phase) bandpass filter of 100–1500 Hz (Corona-Strauss et al., 2009) to the averaged vertex recordings (ABR_{broad}).

For the second analysis, averaged ERPs from the multi-channel recordings were (zero-phase) filtered with the identical 100–1500 Hz bandpass and submitted to a principal component analysis (PCA). This allowed taking advantage of the spatial information provided by the high-density recordings for separating the ERPs into components with varying degrees of contribution to the overall neural activity across the scalp and provided a denoising effect for the previously minimally cleaned data. PCA was carried out over the latency range between -10 ms and 50 ms to maximize sensitivity towards the ABR, and the principal component that accounted for most of the variance in the sensor recordings was identified as the ABR_{PCA}.

For the third analysis, the averaged vertex ERPs were (zero-phase) filtered with the narrow bandpass (105.3–185.0 Hz) that was identified in the ITPC analysis (see Fig. 3, top) to provide significant separation between attended and ignore conditions in the ABR latency range (ABR_{narrow}). It should be noted, however, that the finding of a single VMD mode representing most of the ABR energy does not imply that the different ABR peaks arise from a single neural source.

The fourth analysis examined attention effects on the ABR without any a priori specifications about the spectral content of the intrinsic activity. To this end, averaged vertex ERPs were decomposed into seven intrinsic modes by applying a variational mode decomposition (VMD) using the alternate direction method of multipliers (Dragomiretskiy and Zosso, 2014). As the VMD employs an optimization of the Hardy space representation of the modes in terms of their H^1 -Sobolev regularity for spectral compaction in the Fourier domain, the analysis was again restricted to the latency range between -10 ms and 50 ms to keep the

degree of non-stationarity low. The center frequencies of the modes were initialized as the peaks of each ERP's Fourier spectrum, and the minimization of the augmented Lagrangian functional was performed using a regularization constant of $\alpha = 1000$ and a maximum number of 10 000 iterations. The intrinsic ABR activity was detected by averaging the resulting modes across participants and analyzing them in the frequency-domain. Since only the fourth mode was concentrated well within the ABR-specific passband (105.3–185.0 Hz) with nearly identical peak frequencies of 126.5 Hz for the attended and 126.1 Hz for the ignored condition, it was selected from each individual decomposition as the ABR mode (ABR_{VMD}).

The four ABR representations were reduced to scalar features by extracting the amplitudes of the prominent peaks. These were identified as wave V peaking at 5 ms and subsequent deflections that were labeled according to their (approximate) peak latencies, including N₉, P₁₂, and N₁₆. For each of the peaks, the amplitude was determined as the mean voltage within a time window of -1.5 ms to 1.5 ms relative to peak latency.

2.8. ERP spectrotemporal filtering analysis

The broadband (1–1500 Hz) waveforms of the full-range ERPs (see Fig. 1) only presented significant effects of attention in the AMLR and ALR latencies. However, all subsequent analyses indicated that attention also modulated the ABR when targeting its characteristic time and frequency interval. In order to demonstrate the full range of attention effects in a broadband time-domain representation, a CWT-based spectrotemporal filtering procedure was performed to adaptively enhance the representations of the ABR, AMLR, and ALR in the waveforms of the full-range ERPs. The analysis was implemented by computing the CWT of the averaged ERPs from vertex recordings, weighting the resulting coefficients by a filter mask that followed the spectrotemporal trend of the full-range ERP (see Fig. S3 in the SI), and applying the inverse CWT to the masked coefficients to reconstruct the time-domain ERPs (ERP_{rec}). The CWT was again performed using analytic Morse wavelets and the identical time and frequency ranges as well as wavelet family parameter as used for the power and ITPC analyses described above, but now the number of wavelet cycles had to be fixed across scales to be consistent with the implementation of the inverse CWT. To optimize the sensitivity towards the ABR, its filter mask center frequency was identified (142.6 Hz) and the corresponding number of cycles from the previous analyses (1.67) was chosen to create the wavelet family.

2.9. Statistics

All statistical analyses described below to investigate effects of attention were based on contrasting the attended versus the ignored condition using two-tailed within-participants *t*-tests and *p*-value adjustments for multiple comparisons following the Benjamini–Hochberg false discovery rate (FDR) correction procedure (Benjamini and Hochberg, 1995).

Broadband ERP waveforms (Fig. 1), time–frequency ITPC and power distributions (Fig. 3), and spectrotemporally filtered ERP waveforms (Fig. 6) were analyzed in an exploratory manner. Waveforms were tested at each time point over the interval 0–250 ms, and *p*-values were FDR-corrected across time ($q = 0.05$). Similarly, ITPC and power maps were analyzed at each time–frequency point within the same time interval and over the frequency range 4–512 Hz, and *p*-values were FDR-corrected across the time–frequency plane ($q = 0.05$). To obtain a more spectrotemporally localized representation of the neural activity components contributing to the differences between conditions, FDR correction was also repeated with a more stringent criterion ($q = 0.01$). Furthermore, statistical analysis of ITPC included an additional stage at which effects were only classified as significant if they were located within a certain “area of trust”. Time–frequency points were assigned to “areas of trust” if the mean ITPC across participants indicated a

significant ($p < 0.01$) deviation from a circular random distribution as analyzed via Rayleigh's test (Rao Jammalamadaka and Sengupta, 2001).

For each of the four ABR representations (Figs. 2 and 4), effects of attention on peak amplitudes were individually tested for wave V, N₉, P₁₂, and N₁₆. The statistical details for these contrasts as well as additional statistics for the PCA and VMD analyses are given in the SI.

In the waveform consistency analysis (Fig. 5), the waveforms of single-trials were correlated with the waveforms of the corresponding averages, separately for the narrowly filtered ABR, AMLR, and ALR in their characteristic latency ranges. Mean correlations across sweeps were tested in a channel-wise manner for vertex and multi-channel recordings. The resulting *p*-values for multi-channel recordings were FDR-corrected across the scalp ($q = 0.05$), separately for the ABR, AMLR, and ALR. Detailed statistical results for the vertex recordings are provided in the SI.

2.10. Behavioral data

After participants' debriefing, it was found that 35.5% of the participants preferred to discriminate between standard and target stimuli for the chirps, while 32.3% preferred the tone bursts, 19.3% had no preference, and 12.9% had a better-ear preference. Objective comparisons between task performances for the two types of stimuli were carried out via two different approaches which were both based on 2×2 confusion matrices (true/false positive/negative). The matrices were calculated for each experimental trial, and button presses were only classified as true positive if the responses occurred within 50–1500 ms after target stimuli onsets. For the first analysis, the confusion matrices were used to derive the true and false positive rates, and the values were averaged across trials. The mean true positive rates across participants were $73.7 \pm 15.3\%$ for the chirps and $69.4 \pm 15.2\%$ for the tone bursts, and the associated false positive rates were $0.3 \pm 0.3\%$ and $0.2 \pm 0.2\%$, respectively. These results confirmed that participants were successful in detecting the target stimuli. For the second analysis, the four cells in each contingency table were summarized using the Matthews correlation coefficient (MCC) (Chicco and Jurman, 2020). This measure can take values between -1 (perfect misclassification) and 1 (perfect classification), with a value of 0 indicating a random classification. Importantly, the MCC provides unbiased results for not only balanced but also imbalanced binary classification tasks as opposed to other popular measures such as the F1 score (Chicco and Jurman, 2020), which was particularly important for the present study due to the difference in the numbers of standard and target stimuli. The MCC was calculated for each confusion matrix, and the values were again averaged across trials. While the mean MCC across participants was slightly higher for chirps (0.82 ± 0.10) than for tones (0.80 ± 0.10), a within-participants *t*-test (two-tailed) confirmed that the stimulus type had no significant effect on task performance ($t(30) = 1.36$, $p = 0.18$).

3. Results

Fig. 1 shows the grand-average broadband (1–1500 Hz) waveforms of the vertex ERPs elicited by the tones and the chirps under attended and ignored conditions. The logarithmic time base facilitates the visualization of the full-range ERP including the ABR, AMLR, and ALR (the same figures with a linear time base are in Fig. 2. in the SI). For both types of stimuli, there was a clear enhancement of the N1 component of the ALR (peaking at 90–100 ms) to attended-ear sounds as in the original N1 effect study (Hillyard et al., 1973). As expected, however, the chirps elicited prominent early components in the ABR and AMLR latency range that were absent in the tone-evoked ERPs (see SI section on chirp-evoked ABRs). In particular, the chirp-evoked waveform showed a vertex-positive peak with a latency of around 5 ms, which corresponded to wave V of the ABR (Maloff and Hood, 2014).

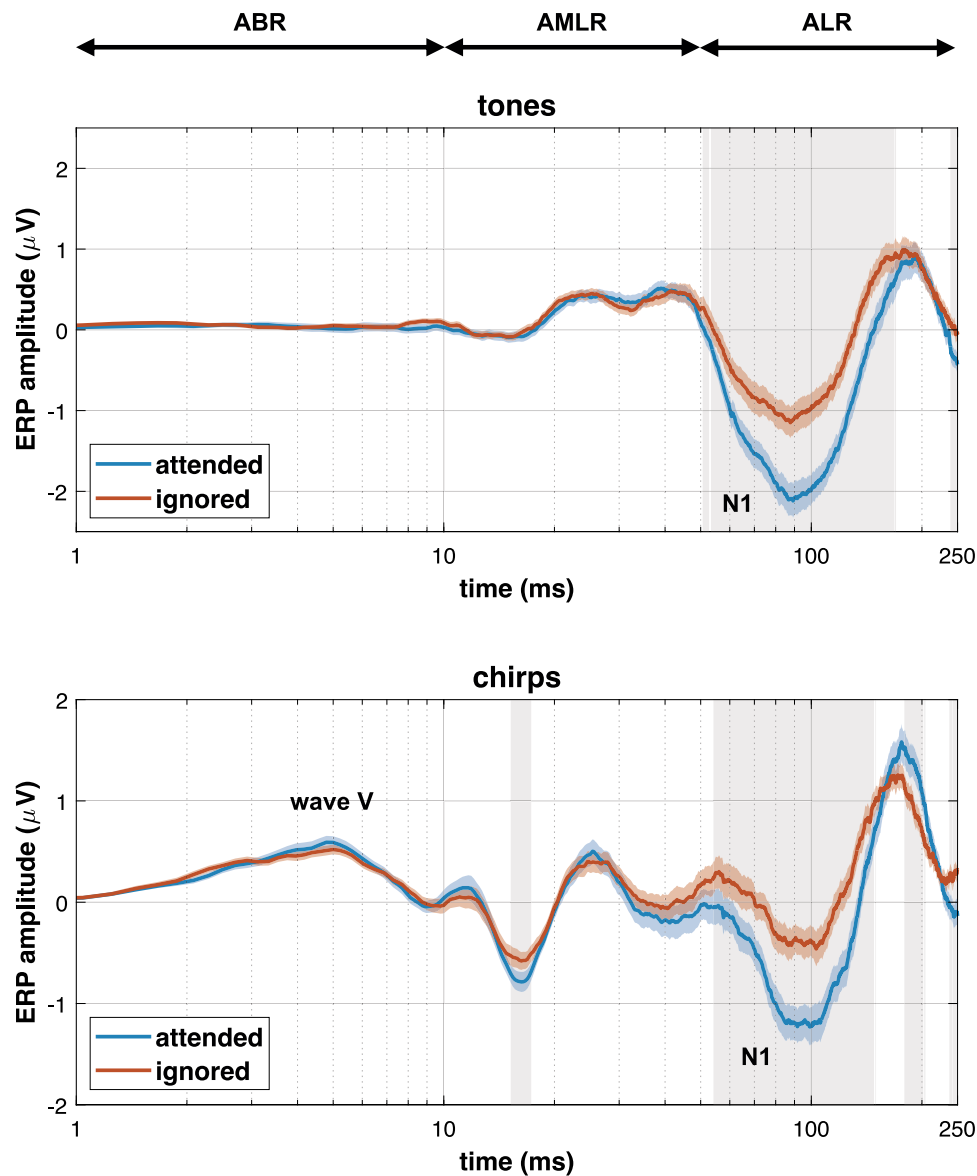


Fig. 1. Grand-average ($N = 31$ participants) broadband (1–1500 Hz) waveforms of vertex ERPs elicited by the tones (top) and the chirps (bottom) on a logarithmic time base. The colored shaded backgrounds show the condition-specific standard errors across participants and gray areas highlight time periods with significant differences between conditions ($p < 0.05$; two-tailed within-participants t -test with Benjamini–Hochberg FDR correction across time). While the ERPs to both types of stimuli showed a prominent N1 effect at around 90–100 ms, the chirp-evoked ERPs exhibit a more complex structure in the ABR and AMLR latency range than the tone-evoked waveforms. This is an expected consequence of the synchronous activation of the cochlea by the chirps.

Since the major focus of this study was the effect of attention on the ABR, subsequent analyses were focused on the chirp-evoked ERPs.

As seen in the broadband ERPs to the chirps (Fig. 1, bottom), wave V was larger in the attended versus ignored grand-average waveform, but this difference did not reach statistical significance; this is most likely due to the broad passband required to represent all the spectrotemporal components of the full-range ERP, which is not optimal for resolving the high-frequency ABR components. Indeed, when the averaged waveforms were filtered with a conventional bandpass for chirp-evoked ABRs (ABR_{broad}) of 100–1500 Hz (Corona-Strauss et al., 2009), the attention effects on wave V and subsequent AMLR components were found to be significant in the vertex ERPs as well as in the ABR components derived from a PCA (ABR_{PCA}) of the multi-channel recordings (Fig. 2; see SI for statistical details of the ABR_{broad} and ABR_{PCA} analyses).

To analyze the frequency-specific signatures of top-down attention in the chirp-evoked ERPs at ABR and subsequent latency ranges, we carried out a time–frequency analysis on the single-trial vertex waveforms by means of CWTs using analytic Morse wavelets (Lilly

and Olhede, 2012). This analysis showed that attention produced an increase in ITPC of evoked activity at multiple spectrotemporal scales of the auditory full-range ERP (Fig. 3, top). While an increased ITPC has been reported previously for the time–frequency range that encompasses the N1 component (Trenado et al., 2009; Low and Strauss, 2011), the present analysis with chirp stimulation revealed that corresponding patterns also exist for the ABR and AMLR components. Even though this analysis was tuned for the best possible resolution by an adjustment of the wavelet cycles regarding Heisenberg’s uncertainty in time–frequency analysis (see Methods), we have supported these findings with a super-resolution analysis (Moca et al., 2021) (see Fig. S3 in the SI), to ensure that the time–frequency energy zones in the ABR range were not due to bled-in energy components from lower frequencies or longer latency components.

Importantly, selective attention significantly enhanced the ITPC in a latency range that encompassed wave V of the chirp-evoked ABR. This increase was significant ($p < 0.01$) over the frequency band 105.3–185.0 Hz, which is in line with fiber tract models of the

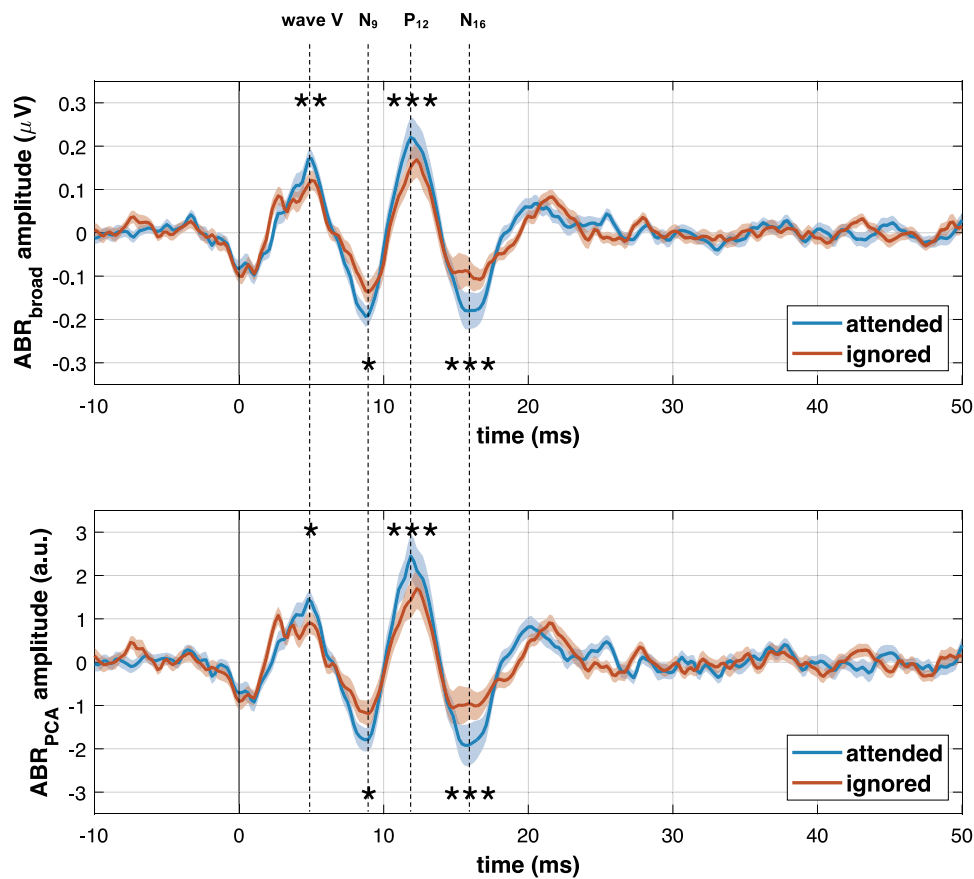


Fig. 2. Grand-average ($N = 31$ participants) waveforms of broadband chirp-evoked ERPs. Time zero corresponds to the end of the chirp stimulus (see Fig. S1 in the SI). ABR_{broad} waveforms (top) were recorded from the vertex and bandpass-filtered within a conventional broad frequency range of 100–1500 Hz. ABR_{PCA} waveforms (bottom) were obtained by applying the same bandpass filter to the averaged multi-channel ERPs and performing a PCA to extract the ABR-specific component. The colored shaded backgrounds show the condition-specific standard errors across participants. Asterisks symbolize statistically significant effects of attention on component peak amplitudes (* $p < 0.05$, ** $p < 0.02$, *** $p < 0.01$; two-tailed within-participants t -test). Although these two ABR representations were obtained from different electrode configurations, they exhibit remarkably similar waveforms and significant effects of attention on all component peak amplitudes including wave V of the ABR as well as subsequent negative and positive peaks within the transitional latency range between ABRs and AMLRs. Statistical details are given in the SI.

ABR (Rudell, 1987). When the averaged vertex ERPs were filtered in this narrow ABR-specific frequency band (ABR_{narrow}), the attentional enhancement of wave V became highly significant but lost some temporal precision in time due to the enlarged impulse response of the associated narrow-band filter (Fig. 4, top). To complement these linear filtering observations with a non-linear decomposition to address time-frequency resolution, ABR representations from the original broadband waveforms were also obtained using a VMD (ABR_{VMD}) (Dragomiretskiy and Zosso, 2014). In line with our previous arguments, the VMD analysis shows that the high-frequency energy in the wave V interval stems from an intrinsic mode of the ABR and not merely a bleed-in energy byproduct of later ERP components because of Heisenberg's uncertainty principle (see Fig. 4, bottom; see SI for statistical details of the ABR_{narrow} and ABR_{VMD} analyses). In fact, the VMD analysis parallels the morphology of ABR_{narrow} but provides better resolution in time. However, as this approach is non-linear with intrinsic modes computed for each participant, the overall morphology of the modes for the two conditions also deviates from the linear transforms. For instance, wave V was relatively smaller wave V for the ignored condition. Since VMD is computed without prior spectral constraints, the overall diminished energy for the ignored condition of the fourth mode in a 7-mode VMD (see Methods) also reflects less spectrotemporal consistency across participants for the unattended stimuli. It is evident that all ABR representations presented in Figs. 2 and 4, even though derived from different EEG montages as well as linear and non-linear approaches, consistently exhibit attention-driven amplitude enhancements in the averaged ERPs for wave V of the ABR and subsequent components.

While attention significantly increased the time-domain amplitude of wave V and subsequent AMLR peaks in several different ABR representations of averaged responses (Figs. 2 and 4) and significantly increased the wavelet ITPC (Fig. 3, top) at each characteristic time-frequency scale of the auditory full-range ERP, the associated increases of wavelet power in the same time-frequency zones did not reach statistical significance (Fig. 3, bottom). This pattern of results could be attributed to the relative robustness of phase synchrony measures using circular statistics (e.g., ITPC) as compared to amplitude (power) measures in single-trial ERPs under conditions of low signal-to-noise ratios (Trenado et al., 2009; Low and Strauss, 2011; Rosenblum et al., 2001; van Diepen and Mazaheri, 2018). However, this data pattern is also consistent with a mechanism whereby attention increases the synchronization as well as the amplitude of evoked neural activity, which would increase the ITPC to a greater extent than the power in the wavelet analysis. To investigate whether attention also increases waveform consistency on a single-trial basis in the time-domain, we calculated the mean correlations of single-trial vertex ABR, AMLR, and ALR waveforms with their corresponding averaged waveforms. The resulting correlations were increased by attention for each sub-interval of the full-range ERP (Fig. 5, top; see SI for statistical details) which, along with the ITPC analysis (Fig. 3, top), demonstrates that selective attention produces a time-locked morphological stability of single-trial neural activity from brainstem up to cortical levels of the auditory pathways. These increases in waveform consistency with attention were evident over widespread areas of the scalp (Fig. 5, bottom) consistent with the well-known broad scalp distributions of auditory ERP components (Picton, 2010).

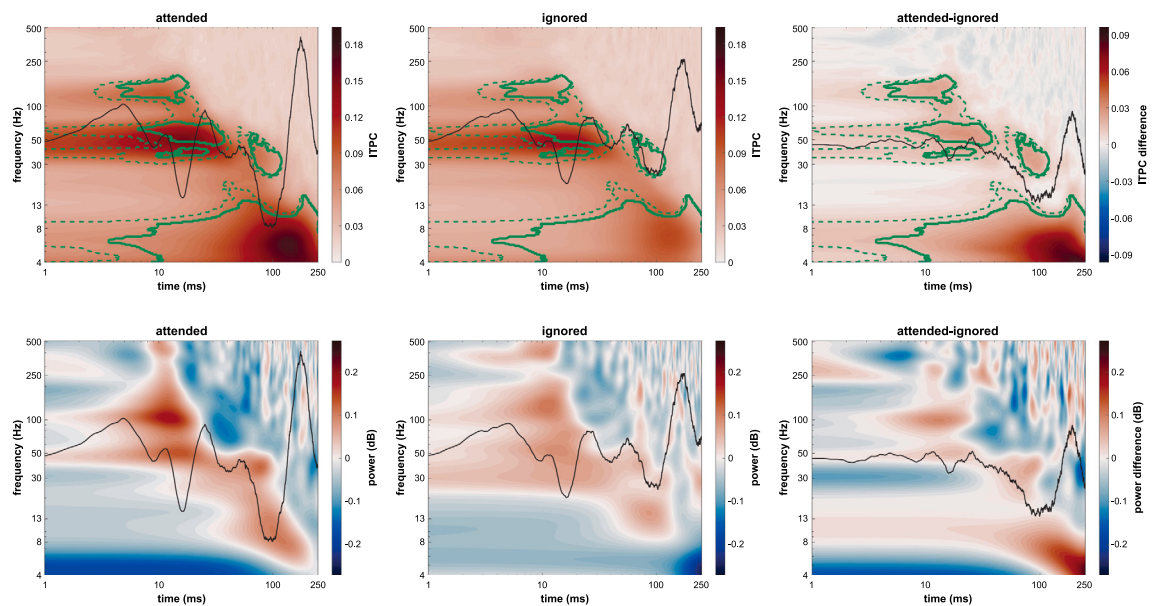


Fig. 3. Grand-average ($N = 31$ participants) time-frequency ITPC (top) and power (bottom) maps for the single-trial wavelet analysis of vertex waveforms for attended (left) and ignored (middle) chirps as well as the resulting difference map (right; attended-ignored). Black waveforms represent identically scaled condition-specific grand-average ERPs (left and middle; same as Fig. 1) and their difference wave (right). Green lines circumscribe time-frequency zones with significant differences between conditions ($p < 0.05$ for dashed lines and $p < 0.01$ for solid lines; two-tailed within-participants t -test with Benjamini–Hochberg FDR correction across time and frequency). While the ITPC was significantly increased by attention at the spectrotemporal scales of the ABR, AMLR, and ALR, the apparent wavelet power modulations in the same time-frequency zones did not reach significance.

The foregoing analyses showed that the effects of attention on wave V and subsequent AMLR components did not reach significance in the time-domain amplitude of the averaged broadband ERPs (Fig. 1, bottom) but did become significant in several ABR-specific waveform representations (Figs. 2 and 4). Critically, the wavelet ITPC analysis (Fig. 3, top) revealed that higher spectrotemporal consistency across single-trials for attended versus ignored stimuli ($p < 0.01$) was specific to time–frequency ranges matching the scales of the ABR, AMLR, and ALR, which indicated that the non-specific filtering used for the analysis shown in Fig. 1 may have masked putative attention effects in those broadband waveforms. Accordingly, to analyze the attentional modulations in the time-domain amplitude of the averaged full-range ERP more precisely, we implemented a CWT-based spectrotemporal filtering approach to enhance the representation of the individual components within their corresponding latency ranges. In brief, averaged chirp-evoked ERPs recorded from the vertex were mapped to time–frequency representations via the CWT using analytic Morse wavelets, the wavelet coefficients were weighted with a filter mask (see Fig. S4 in the SI) that was based on the ITPC significance patterns, and full-range ERPs were reconstructed via the inverse CWT of the weighted coefficients. The resulting ERP reconstructions (ERP_{rec}) for both conditions are shown in Fig. 6. While the spectrotemporally filtered ERPs exhibited less noise compared to the original broadband ERPs (Fig. 1, bottom), their morphologies were almost identical, with clearly matching positive and negative peaks. Importantly, however, the localized filtering in time and frequency revealed significant effects of attention on the time-domain amplitude at each temporal scale of the auditory full-range ERP (wave V to N1) that were previously obscured due to the standard broadband filtering that was applied in Fig. 1.

4. Discussion

A classical question in both neural and behavioral studies of auditory selective attention has been whether and under what circumstances the preferential selection of attended stimuli occurs at early or late stages of processing (Hernández-Peón et al., 1956; Treisman and Geffen, 1967).

The present study investigated whether attention to brief chirp stimuli, optimized to stimulate the entire cochlea synchronously, would reveal modulation of neural transmission in the human auditory brain-stem pathways. In a dichotic listening paradigm similar to that used by Hillyard et al. (1973), chirps were presented to one ear and tones to the other in a rapid, unpredictable sequence to prevent different levels of arousal or other non-selective preparatory states from arising prior to the attended stimuli. The key finding was that when chirp stimuli were attended as compared to ignored, the elicited ERPs exhibited an increased phase consistency at the time–frequency scale of wave V of the ABR as well as a corresponding enhancement in time-domain amplitude for the averaged potentials. The generation of this component has been anatomically linked to evoked activity in the IC (Møller and Jannetta, 1985; Yvert et al., 2002). The attentional modulation of wave V took place as early as 5 ms after chirp presentation and was followed by enhanced amplitudes of multiple components in the AMLR (10–50 ms) and ALR (50–250 ms) latency ranges in the averaged waveforms. These results provide critical evidence for very early stimulus selection processes in human auditory attention, at least under the present conditions of dichotic listening using brief tonal stimuli.

Previous studies of attention effects on ABRs have not yielded consistent results. Most studies reported no effect of selective attention on ABR amplitudes (Picton et al., 1971, 1974; Woods and Hillyard, 1978; Hackley et al., 1987, 1990; Collet and Duclaux, 1986; Connolly et al., 1989; Gregory et al., 1989; Hirschhorn and Michie, 1990; Woldorff and Hillyard, 1991), and the few suggestive effects (Lukas, 1980, 1981) have been challenged (Hirschhorn and Michie, 1990; Hoormann et al., 2000) or were of marginal significance (Kumar et al., 2023). The present study differed from its predecessors by using chirp stimuli rather than clicks or tone pips. It seems reasonable to propose that the wave V enhancement observed here was a consequence of the synchronous activation of the entire cochlea and afferent pathways produced by the chirps, which allowed the descending attentional control to be manifest. The physiological mechanism of this attention effect on wave V remains unknown, but it may be mediated by the massive efferent cortico-collicular projections, which have been shown in animal studies to modulate synaptic activity in the IC (Suga et al.,

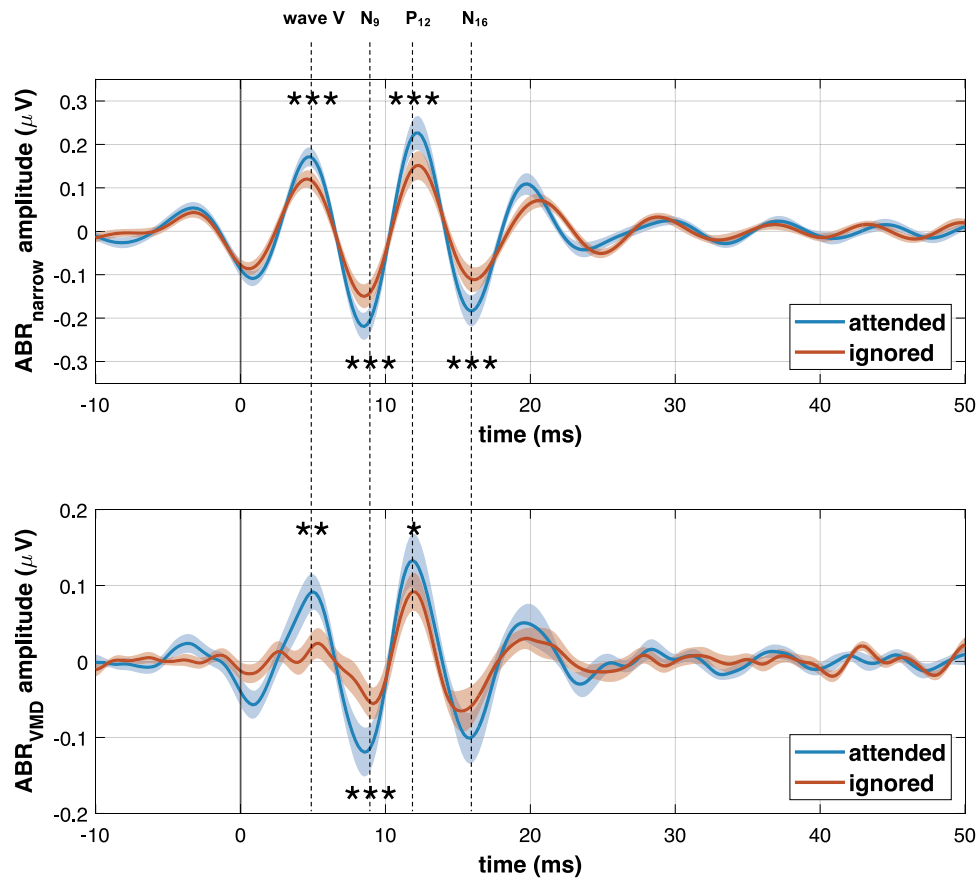


Fig. 4. Grand-average ($N = 31$ participants) waveforms of enhanced narrowband ABR representations from chirp-evoked ERPs. ABR_{narrow} waveforms (top) were extracted via bandpass filtering of averaged vertex waveforms within the ABR frequency range of 105.3–185.0 Hz that showed significant ($p < 0.01$) increases in ITPC with attention. ABR_{VMD} waveforms (bottom) were obtained by performing a VMD on the same data and identifying the ABR-specific mode. The colored shaded backgrounds show the condition-specific standard errors across participants. Asterisks symbolize statistically significant effects of attention on component peak amplitudes (* $p < 0.05$, ** $p < 0.02$, *** $p < 0.01$; two-tailed within-participants t -test). The non-linear VMD procedure yielded waveforms that were highly similar to those based on linear filtering in the identified ABR-specific frequency range. Importantly, both narrowband ABR representations exhibited attentional modulations of wave V and subsequent component peaks that are fully congruent with the effects observed for ABR_{broad} and ABR_{PCA} (Fig. 2). Statistical details are given in the SI.

2002; Blackwell et al., 2020; Suga and Ma, 2003; Oberle et al., 2023). It is also possible, however, that attentional control could be exerted at even earlier levels of the auditory pathways and passed along to the IC. For example, there have been reports that stimulus-evoked otoacoustic emissions (EOAEs) arising from the cochlea can be modulated by attention (Giard et al., 1994; Saiz-Alfía et al., 2021), which would implicate descending attentional control via the olivo-cochlear bundle. Attention effects on the EOAE in humans, however, have not been consistently observed (Michie et al., 1996).

Previous studies using differing methodologies have also reported attentional control over neural activity in the auditory brainstem. In a functional magnetic resonance imaging (fMRI) study, Rinne et al. (2008) presented sequences of noise bursts to the left and right ears and found that activation in the IC depended on the direction of attention. While these fMRI measures provide precise localization of the neural modulation with attention, they do not give information about the timing of the attention effects nor whether they specifically reflect attentional modulation of sensory evoked afferent activity in the IC. In a series of EEG studies of the auditory FFR to continuous speech, Reichenbach and colleagues extracted brainstem response waveforms at the fundamental frequency (termed “speech-ABRs”) by means of empirical mode decomposition (Forte et al., 2017; Etard et al., 2019; Saiz-Alfía et al., 2019). The speech-ABR had a lag of about 8 ms relative to the acoustic speech waveform and was found to be enlarged when a spoken message was attended in a dichotic listening situation. Computational modeling of the subcortical origins of the speech-ABR identified the IC as its dominant source (Saiz-Alfía and Reichenbach, 2020). It is not

clear, however, how rapidly attentional selectivity of the FFR began after sound onset or whether the attentional control mechanisms over continuous sinusoidal inputs such as speech are the same as those engaged in attending to singular stimulus onsets (Bidelman, 2015). The present finding of attentional modulation of wave V of the ABR to an individual, abruptly onsetting stimulus is thus of critical importance for understanding how selective attention is able to modulate afferent signals in the auditory brainstem.

In addition to the ABR modulations, attending to the chirps resulted in enlarged amplitudes of later AMLR and ALR components in the time-domain averages, including the iconic N1 enhancement. The chirp-evoked AMLR had a more complex, multi-component structure than the tone-evoked AMLR, which can be ascribed to the abrupt and synchronous activation of the cochlea (and higher pathways) by the chirps. Previous studies reported an enhanced positivity in the 20–50 ms range to attended high-frequency tone pips, which were localized to auditory cortex by MEG recordings (Woldorff and Hillyard, 1991; Woldorff et al., 1993). This P20–50 effect may correspond to the enhanced positivities at around 25 ms and 40–60 ms observed here. Dipole modeling studies by Scherg and Von Cramon (Scherg and Von Cramon, 1985, 1986) determined that the major components of both the AMLR and ALR represented evoked activity in the auditory cortex. The present results thus demonstrate that in this dichotic listening situation afferent signals are amplified by selective attention all along the auditory pathway from brainstem to cortex.

The effects of attention on evoked activity in the brainstem and higher levels of the auditory pathways were evident not only in the

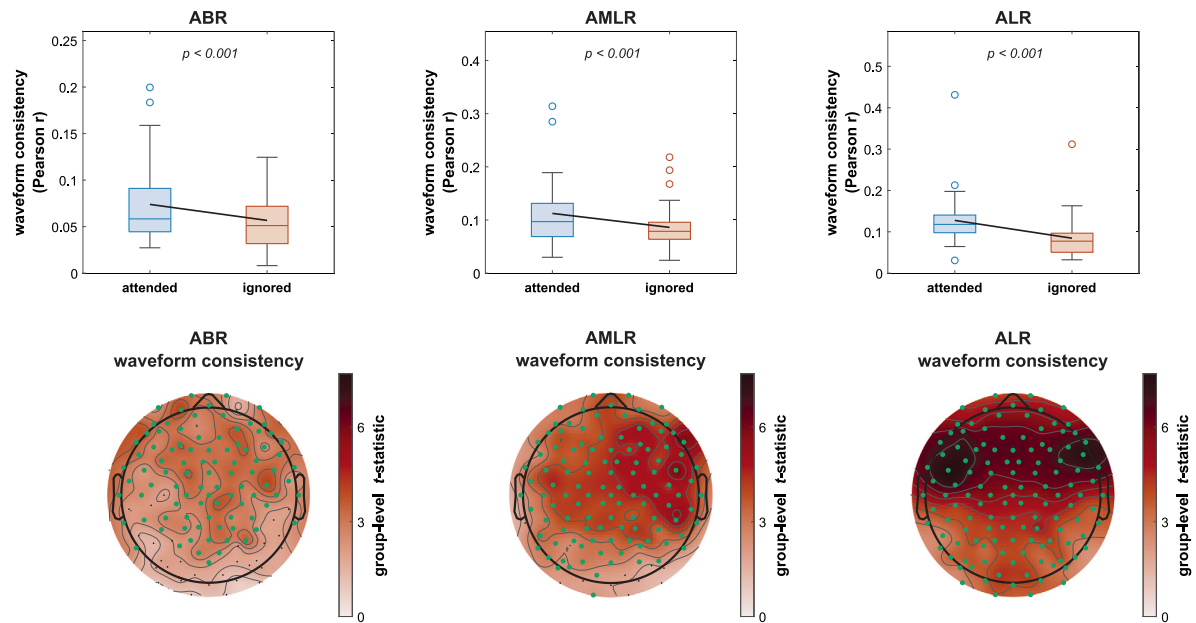


Fig. 5. Waveform consistency analysis for the chirp-evoked ERPs. Consistency was calculated as the mean Pearson correlation between the averaged ERP and its contributing single-trials for narrowband representations of the ABR (0–10 ms, 105.3–185.0 Hz; left), AMLR (10–50 ms, 44.3–71.3 Hz; middle), and ALR (50–250 ms, 4.0–15.7 Hz; right). The figure presents the box plots ($N = 31$ participants) for the mean correlation values for vertex recordings (top) and the group-level t -statistic topographies for multi-channel recordings (bottom; attended versus ignored correlations). Channels that indicated significant differences between conditions are marked by green dots. For each sub-interval of the auditory full-range ERP, attention increased the consistency in vertex recordings (all $p < 0.001$; two-tailed within-participants t -test) as well as for the vast majority of electrodes in multi-channel recordings ($p < 0.05$; two-tailed within-participants t -test with Benjamini–Hochberg FDR correction across channels). Statistical details for the vertex recordings are given in the SI.

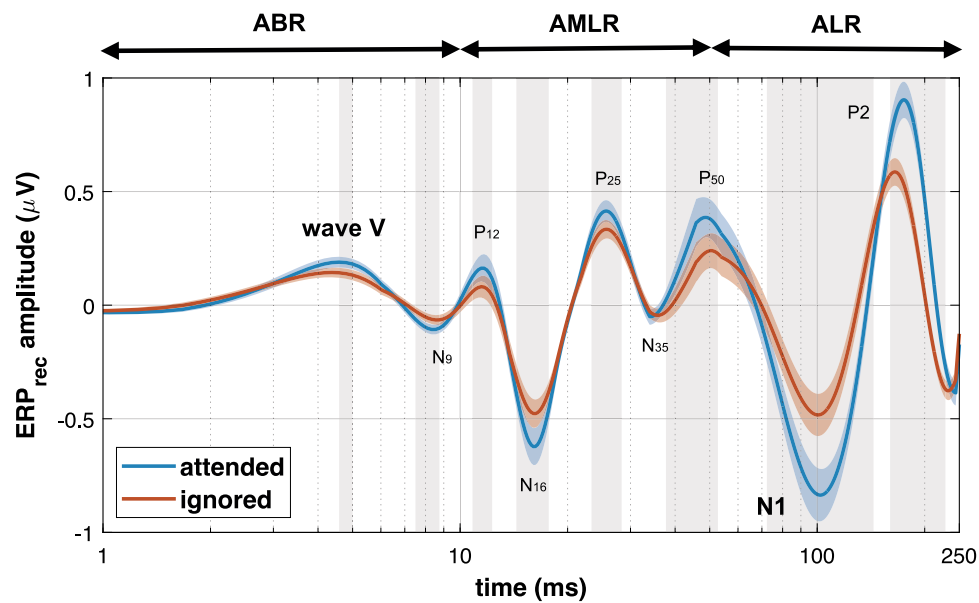


Fig. 6. Grand-average ($N = 31$ participants) waveforms of spectrotemporally filtered reconstructions of averaged chirp-evoked ERPs from vertex recordings. Time–frequency representations of the ERPs were weighted by a filter mask (see Fig. S4 in the SI) and back-transformed to the time-domain. The colored shaded backgrounds show the condition-specific standard errors across participants and gray areas highlight time periods with significant differences between conditions ($p < 0.05$; two-tailed within-participants t -test with Benjamini–Hochberg FDR correction across time). While the waveform morphologies closely paralleled the ones of the original broadband ERPs (Fig. 1, bottom), the adaptive filtering approach revealed significant modulations with attention at each sub-interval of the full-range response, even at early brainstem levels such as wave V of the ABR.

conventional time-domain averages but also in the wavelet-based ITPC and the correlations of single-trial waveforms with the overall averages. The link between amplitude variations and the instantaneous phase of ERPs has been well documented (Mortezapouraghdam et al., 2018; Benhamou et al., 2023), which suggests that the ERP amplitude increases with attention observed here may be based in part on an enhancement of temporal synchronization of the evoked neural activity to the attended stimuli as expressed in the ITPC as well as the time-domain

correlation analysis. Previous studies have already linked the N1 attention effect to spectrotemporal phase consistency across trials (Trenado et al., 2009; Low and Strauss, 2011), and such phase-locking to the stimulus could also play a similar role in attentional modulation of earlier ABR and AMLR components. If top-down attentional control does produce an increased synchrony of firing in afferent neural populations, this could not only enhance the signal-to-noise ratio in favor of attended signals but also facilitate the integration of different auditory features

into perceived objects (Asilador and Llano, 2021). Further studies are needed to reveal the underlying neurophysiological mechanisms responsible for these facilitatory effects of selective attention on the ABR, AMLR, and ALR in a wider range of task situations and using different types of more complex stimuli. Of particular interest, the effects of task difficulty and varying levels of attentional effort in listening (see Sarter et al. (2006), Strauss and Francis (2017)) on the N1 wave in ERPs have been well-documented (Strauss et al., 2010; Bernarding et al., 2013). Building upon the present findings, studies incorporating such concepts related to different levels of attentional effort and their influence on chirp-evoked full-range potentials represents a promising avenue for future research.

CRediT authorship contribution statement

Daniel J. Strauss: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Farah I. Corona-Strauss:** Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Adrian Mai:** Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Steven A. Hillyard:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors were partially supported by the European Regional Development Fund (ERDF) and the state of Saarland by the “Center for Digital Neurotechnologies Saar (CDNS)” (Projekt-ID: EFRE-HS-0000835).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2025.121295>.

Data availability

The data and source code can be downloaded from neuroimage-25-121295.snn-unit.de.

References

- Asilador, A., Llano, D.A., 2021. Top-down inference in the auditory system: Potential roles for corticofugal projections. *Front. Neural Circuits* 14, 615259.
- Benhamou, J., Le Van Quyen, M., Marrelec, G., 2023. Time-frequency analysis of event-related brain recordings: Connecting power of evoked potential and inter-trial coherence. *IEEE Trans. Biomed. Eng.* 70, 1599–1610.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57, 289–300.
- Bernarding, C., Strauss, D.J., Hannemann, R., Seidler, H., Corona-Strauss, F.I., 2013. Neural correlates of listening effort related factors: Influence of age and hearing impairment. *Brain Res. Bull.* 91, 21–30.
- Bidelman, G.M., 2015. Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. *Hear. Res.* 323, 68–80.
- Blackwell, J.M., Lesicko, A.M., Rao, W., De Biasi, M., Geffen, M.N., 2020. Auditory cortex shapes sound responses in the inferior colliculus. *eLife* 9, e51890.
- Chicco, D., Jurman, G., 2020. The advantages of the matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21, 6.
- Coffey, Emily B.J., Herholz, Sibylle C., Chepesiuk, Alexander M.P., Baillet, Sylvain, Zatorre, Robert J., 2016. Cortical contributions to the auditory frequency-following response revealed by meg. *Nat. Commun.* 7, 11070.
- Collet, L., Duclaux, R., 1986. Auditory brainstem evoked responses and attention. Contribution to a controversial subject. *Acta Otolaryngol.* 101, 439–441.
- Connolly, J.F., Aubry, K., McGillivray, N., Scott, D.W., 1989. Human brainstem auditory evoked potentials fail to provide evidence of efferent modulation of auditory input during attentional tasks. *Psychophysiol.* 26, 292–303.
- Corona-Strauss, F.I., Delb, W., Schick, B., Strauss, D.J., 2009. Phase stability analysis of chirp evoked auditory brainstem responses by gabor frame operators. *IEEE Trans. Neural Syst. Rehabil. Eng.* 17, 530–536.
- Dau, T., Wegner, O., Mellert, V., Kollmeier, B., 2000. Auditory brainstem responses (ABR) with optimized chirp signals compensating basilar-membrane dispersion. *J. Acoust. Soc. Am.* 107, 1530–1540.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.
- van Diepen, R.M., Mazaheri, A., 2018. The caveats of observing inter-trial phase-coherence in cognitive neuroscience. *Sci. Rep.* 8, 2990.
- Dragomiretskiy, K., Zosso, D., 2014. Variational mode decomposition. *IEEE Trans. Signal Process.* 62, 531–544.
- Elberling, C., Callø, J., Don, M., 2010. Evaluating auditory brainstem responses to different chirp stimuli at three levels of stimulation. *J. Acoust. Soc. Am.* 128, 215–223.
- Elberling, C., Don, M., 2008. Auditory brainstem responses to a chirp stimulus designed from derived-band latencies in normal-hearing subjects. *J. Acoust. Soc. Am.* 124, 3022–3037.
- Etard, O., Kegler, M., Braiman, C., Forte, A.E., Reichenbach, T., 2019. Decoding of selective attention to continuous speech from the human auditory brainstem response. *NeuroImage* 200, 1–11.
- Fobel, O., Dau, T., 2004. Searching for the optimal stimulus eliciting auditory brainstem responses in humans. *J. Acoust. Soc. Am.* 116, 2213–2222.
- Forte, A.E., Etard, O., Reichenbach, T., 2017. The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *eLife* 6, e27203.
- Galbraith, G.C., Arroyo, C., 1993. Selective attention and brainstem frequency-following responses. *Biol. Psychol.* 37, 3–22.
- Galbraith, G.C., Bhuta, S.M., Choate, A.K., Kitahara, J.M., Mullen, Jr., T.A., 1998. Brain stem frequency-following response to dichotic vowels during attention. *NeuroReport* 9, 1889–1893.
- Giard, M.-H., Collet, L., Bouchet, P., Pernier, J., 1994. Auditory selective attention in the human cochlea. *Brain Res.* 633, 353–356.
- Giard, M.H., Fort, A., Mouchetant-Rostaing, Y., Pernier, J., 2000. Neurophysiological mechanisms of auditory selective attention in humans. *Front. Biosci.* 5, 84–94.
- Gregory, S.D., Heath, J.A., Rosenberg, M.E., 1989. Does selective attention influence the brain-stem auditory evoked potential? *Electroencephalogr. Clin. Neurophysiol.* 73, 557–560.
- Hackley, S.A., Woldorff, M., Hillyard, S.A., 1987. Combined use of microreflexes and event-related brain potentials as measures of auditory selective attention. *Psychophysiol.* 24, 632–647.
- Hackley, S.A., Woldorff, M., Hillyard, S.A., 1990. Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. *Psychophysiol.* 27, 195–208.
- Hansen, J.C., Hillyard, S.A., 1980. Endogenous brain potentials associated with selective auditory attention. *Electroencephalogr. Clin. Neurophysiol.* 49, 277–290.
- Hernández-Peón, R., Scherrer, H., Jouvet, M., 1956. Modification of electric activity in cochlear nucleus during attention in unanesthetized cats. *Sci.* 123, 331–332.
- Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. *Sci.* 182, 177–180.
- Hirschhorn, T.N., Michie, P.T., 1990. Brainstem auditory evoked potentials (BAEPs) and selective attention revisited. *Psychophysiol.* 27, 495–512.
- Hoormann, J., Falkenstein, M., Hohnsbein, J., 2000. Early attention effects in human auditory-evoked potentials. *Psychophysiol.* 37, 29–42.
- Kohl, M.C., Schebsdat, E., Schneider, E.N., Niehl, A., Strauss, D.J., Özdamar, Ö., Bohórquez, J., 2019. Fast acquisition of full-range auditory event-related potentials using an interleaved deconvolution approach. *J. Acoust. Soc. Am.* 145, 540.
- Kumar, S., Nayak, S., Pitchai Muthu, A.N., 2023. Effect of selective attention on auditory brainstem response. *Hear. Balance Commun.* 21, 139–147.
- Lehmann, A., Schönwiesner, M., 2014. Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PLoS One* 9, e85442.
- Lilly, J.M., Olhede, S.C., 2012. Generalized morse wavelets as a superfamily of analytic wavelets. *IEEE Trans. Signal Process.* 60, 6036–6041.
- Low, Y.F., Strauss, D.J., 2011. A performance study of the wavelet-phase stability (WPS) in auditory selective attention. *Brain Res. Bull.* 86, 110–117.
- Lukas, J.H., 1980. Human auditory attention: The olivocochlear bundle may function as a peripheral filter. *Psychophysiol.* 17, 444–452.
- Lukas, J.H., 1981. The role of efferent inhibition in human auditory attention: An examination of the auditory brainstem potentials. *Int. J. Neurosci.* 12, 137–145.

- Maloff, E.S., Hood, L.J., 2014. A comparison of auditory brain stem responses elicited by click and chirp stimuli in adults with normal hearing and sensory hearing loss. *Ear Hear.* 35, 271–282.
- Michie, P.T., LePage, E.L., Solowij, N., Haller, M., Terry, L., 1996. Evoked otoacoustic emissions and auditory selective attention. *Hear. Res.* 98, 54–67.
- Moca, V.V., Bărzan, H., Nagy-Dăbâcan, A., Muresan, R.C., 2021. Time-frequency super-resolution with superlets. *Nat. Commun.* 12, 337.
- Møller, A.R., Jannetta, P.J., 1985. Neural generators of the auditory brainstem response. In: Jacobson, J.T. (Ed.), *The Auditory Brainstem Response*. pp. 13–31.
- Morales, S., Bowers, M.E., 2022. Time-frequency analysis methods and their application in developmental EEG data. *Dev. Cogn. Neurosci.* 54, 101067.
- Mortezapouraghdam, Z., Corona-Strauss, F.I., Takahashi, T., Strauss, D.J., 2018. Reducing the effect of spurious phase variations in neural oscillatory signals. *Front. Comput. Neurosci.* 12, 82.
- Näätänen, R., 1967. Selective attention and evoked potentials. *Ann. Acad. Sci. Fenn. B* 151, 1–226.
- Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42, 313–329.
- Näätänen, R., Picton, T.W., 1987. The N1 wave of the human electric and magnetic response to sound: a review & analysis of the component structure. *Psychophysiol.* 24, 375–425.
- Oatman, L.C., Anderson, B.W., 1977. Effects of visual attention on tone burst evoked auditory potentials. *Exp. Neurol.* 57, 200–211.
- Oberle, H.M., Ford, A.N., Czarny, J.E., Rogalla, M.M., Apostolides, P.F., 2023. Recurrent circuits amplify corticofugal signals and drive feedforward inhibition in the inferior colliculus. *J. Neurosci.* 43, 5642–5655.
- Picton, T.W., 2010. *Human Auditory Evoked Potentials*. Plural Publishing Inc, San Diego, CA, USA.
- Picton, T.W., Hillyard, S.A., Galambos, R., Schiff, M., 1971. Human auditory attention: a central or peripheral process? *Sci.* 173, 351–354.
- Picton, T.W., Hillyard, S.A., Krausz, H.I., Galambos, R., 1974. Human auditory evoked potentials. *Audiol. Neurootology* 36, 179–190.
- Price, C.N., Bidelman, G.M., 2021. Attention reinforces human corticofugal system to aid speech perception in noise. *NeuroImage* 235, 118014.
- Rao Jammalamadaka, S., Sengupta, A., 2001. *Topics in Circular Statistics*. World Scientific Pub Co Inc.
- Rinne, T., Balk, M.H., Koistinen, S., Autti, T., Alho, K., Sams, M., 2008. Auditory selective attention modulates activation of human inferior colliculus. *J. Neurophysiol.* 100, 3323–3327.
- Rosenblum, M., Pikovsky, A., Kurths, J., Schäfer, C., Tass, P.A., 2001. Phase synchronization: from theory to data analysis. In: Hoff, A.J. (Ed.), *Handbook of Biological Physics*. pp. 279–321.
- Rudell, A.P., 1987. A fiber tract model of auditory brain-stem responses. *Electroencephalogr. Clin. Neurophysiol.* 67, 53–62.
- Saiz-Alfá, A.E., Forte, M., Reichenbach, T., 2019. Individual differences in the attentional modulation of the human auditory brainstem response to speech inform on speech-in-noise deficits. *Sci. Rep.* 9, 14131.
- Saiz-Alfá, M., Miller, P., Reichenbach, T., 2021. Otoacoustic emissions evoked by the time-varying harmonic structure of speech. *eNeuro* 8, ENEURO.0428–20.2021.
- Saiz-Alfá, M., Reichenbach, T., 2020. Computational modeling of the auditory brainstem response to continuous speech. *J. Neural Eng.* 17, 036035.
- Sarter, M., Gehring, W.J., Kozak, R., 2006. More attention must be paid: The neurobiology of attentional effort. *Brain Res. Rev.* 51, 145–160.
- Scherg, M., Von Cramon, D., 1985. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalogr. Clin. Neurophysiol.* 62, 32–44.
- Scherg, M., Von Cramon, D., 1986. Evoked dipole source potentials of the human auditory cortex. *Electroencephalogr. Clin. Neurophysiol.* 65, 344–360.
- Schüller, A., Schilling, A., Krauss, P., Rampp, S., Reichenbach, T., 2023. Attentional modulation of the cortical contribution to the frequency-following response evoked by continuous speech. *J. Neurosci.* 43, 7429–7440.
- Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr. Clin. Neurophysiol.* 38, 387–401.
- Strauss, D.J., Corona-Strauss, F.I., Schroeder, A., Flotho, P., Hannemann, R., Hackley, S.A., 2020. Vestibial auriculomotor activity indicates the direction of auditory attention in humans. *eLife* 9, e54536.
- Strauss, D.J., Corona-Strauss, F.I., Trenado, C., Bernarding, C., Reith, W., Latzel, M., Froehlich, M., 2010. Electrophysiological correlates of listening effort: Neurodynamical modeling and measurement. *Cogn. Neurodynamics* 4, 119–131.
- Strauss, D.J., Francis, A.J., 2017. Toward taxonomic model of attention in effortful listening. *Cogn. Affect. Behav. Neurosci.* 17, 809–825.
- Suga, N., Ma, X., 2003. Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Rev. Neurosci.* 4, 783–794.
- Suga, N., Xiao, Z., Ma, X., Ji, W., 2002. Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 36, 9–18.
- Treisman, A., Geffen, G., 1967. Selective attention: perception or response? *Q. J. Exp. Psychol.* 19, 1–17.
- Trenado, C., Haab, L., Strauss, D.J., 2009. Corticothalamic feedback dynamics for neural correlates of auditory selective attention. *IEEE Trans. Neural Syst. Rehabil. Eng.* 17, 46–52.
- Winer, J.A., Larue, D.T., Diehl, J.J., Hefti, B.J., 1998. Auditory cortical projections to the cat inferior colliculus. *J. Comp. Neurol.* 400, 147–174.
- Woldorff, M.G., Gallen, C.C., Hampson, S.A., Hillyard, S.A., Pantev, C., Sobel, D., Bloom, F.E., 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc. Natl. Acad. Sci.* 90, 8722–8726.
- Woldorff, M.G., Hillyard, S.A., 1991. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr. Clin. Neurophysiol.* 79, 170–191.
- Woods, D.L., Hillyard, S.A., 1978. Attention at the cocktail party: Brainstem evoked potentials reveal no peripheral gating. In: Otto, D.A. (Ed.), *Multidisciplinary Perspectives in Event-Related Brain Potential Research*. pp. 230–233.
- Yvert, B., Fischer, C., Guénot, M., Krolak-Salmon, P., Isnard, J., Pernier, J., 2002. Simultaneous intracerebral EEG recordings of early auditory thalamic and cortical activity in human. *Eur. J. Neurosci.* 16, 1146–1150.