Methodological considerations when measuring and analyzing auditory steady-state responses with multi-channel EEG

Hao Lu [*], Anahita H. Mehta , Andrew J. Oxenham

Department of Psychology, University of Minnesota, 75 East River Parkway, Minneapolis, MN, 55455, USA

ARTICLE INFO

ABSTRACT

The auditory steady-state response (ASSR) has been traditionally recorded with few electrodes and is often measured as the voltage difference between mastoid and vertex electrodes (vertical montage). As high-density EEG recording systems have gained popularity, multi-channel analysis methods have been developed to integrate the ASSR signal across channels. The phases of ASSR across electrodes can be affected by factors including the stimulus modulation rate and re-referencing strategy, which will in turn affect the estimated ASSR strength. To explore the relationship between the classical vertical-montage ASSR and whole-scalp ASSR, we applied these two techniques to the same data to estimate the strength of ASSRs evoked by tones with sinusoidal amplitude modulation rates of around 40, 100, and 200 Hz. The whole-scalp methods evaluated in our study, with either linked-mastoid or common-average reference, included ones that assume equal phase across all channels, as well as ones that allow for different phase relationships. The performance of simple averaging was compared to that of more complex methods involving principal component analysis. Overall, the root-mean-square of the phase locking values (PLVs) across all channels provided the most efficient method to detect ASSR across the range of modulation rates tested here.

1. Introduction

Periodic auditory stimuli, such as pure tones, amplitude-modulated (AM) tones, and voiced speech, can evoke a periodic brain response at the frequency corresponding to the stimulus periodicity, which can be recorded via electroencephalography (EEG) (Aiken and Picton, 2008; Batra et al., 1986; Mehta et al., 2021) or magnetoencephalography (MEG) (Coffey et al., 2016b; Gorina-Careta et al., 2021; Kulasingham et al., 2020); for a review, see Krizman and Kraus (2019). The periodic response to the temporal fine structure of pure tones is known as the frequency-following response (FFR) (Batra et al., 1986; Gerken et al., 1975; Glaser et al., 1976). For complex stimuli, where the neural response is to the repetition rate of the temporal envelope, it has been termed the auditory steady-state response (ASSR) (Rance, 2008) or the envelope-following response (EFR) (Holmes et al., 2018), often interchangeably. Although we reserve the term FFR for neural responses to the temporal fine structure of pure tones, this term has also been used in the literature to refer to temporal-envelope responses or responses that reflect both temporal fine structure and envelope (Coffey et al., 2019). To avoid such potential confusion, we use the more general term of ASSR

here. The ASSR has been used in clinical tests to provide an objective estimate of hearing thresholds across a range of audiometric frequencies (Duarte et al., 2008; Vander Werff, 2009) and has been employed in auditory research studies to separate the neural representation of different aspects of complex sounds by, for instance, selectively applying AM to only some tones within a complex mixture (Bharadwaj et al., 2014; Gutschalk et al., 2008; Heo et al., 2017) or applying different AM rates to different carriers (Gander et al., 2010; Lau et al., 2017; Mehta et al., 2021; Polonenko and Maddox, 2019).

The ASSR recorded with EEG is the summation of phase-locked activity from multiple neural generators within the auditory system, including the cochlea, auditory nerve, inferior colliculus, and auditory cortex. One important factor affecting the relative contributions of these sources is the rate at which the stimulus is modulated (Coffey et al., 2019; Gnanateja et al., 2021; Krizman and Kraus, 2019; Tan et al., 2015). In general, for EEG studies the contribution of subcortical sources, relative to that of cortical sources, is thought to increase with increasing modulation rate, with the ASSR in response to lower ($<$40 Hz) rates dominated by cortical components (Aiken and Picton, 2008; Herdman et al., 2002) and the ASSR in response to higher ($>$100 Hz)

rates dominated by subcortical components (Bidelman, 2018; Chandrasekaran and Kraus, 2010; Herdman et al., 2002). For all modulation rates there will be some contribution from peripheral sources, e.g., cochlear microphonic (Sohmer et al., 1977) and the auditory nerve (Bidelman, 2018), but their relative contribution to the ASSR depends on the applied measurement technique (Galbraith et al., 2000). Additionally, the ASSR recorded with EEG usually requires a large number of trials due to its low signal-to-noise ratio, especially for situations with more subcortical components (Krizman and Kraus, 2019; Skoe and Kraus, 2010).

Selecting an appropriate EEG electrode configuration or montage is an important part of the experimental design. The ASSR at higher frequencies (≥100 Hz) has been traditionally recorded via a vertical forehead-to-neck electrode montage consisting of a fronto-central non-inverting vertex electrode and an inverting reference electrode, such as left or right mastoid reference (M1 or M2), linked mastoids reference (LMR), or earlobe reference (Batra et al., 1986; Gerken et al., 1975; Greenberg et al., 1987; Krizman and Kraus, 2019; Skoe and Kraus, 2010). This vertical montage remains in common use (Billings et al., 2019; Cone-Wesson et al., 2002; Krishnan et al., 2004; Tichko and Skoe, 2017), and the amplitude of the resulting ASSR has been interpreted as a measure of the strength of an assumed vertical dipole in the auditory brainstem (Galbraith, 1994; Krizman and Kraus, 2019). The existence of this dominant dipole has been confirmed both by source-localization studies using high-density EEG (Bidelman, 2015, 2018; Chandrasekaran and Kraus, 2010; Coffey et al., 2019) and by a double-dissociation human case study (White-Schwoch et al., 2019). Although different choices of the inverting reference channel (M1, M2, or LMR) do not appear to affect the ASSR (Galbraith, 1994), selecting different non-inverting vertex channels can lead to ASSRs that emphasize different aspects of subcortical activity (Parthasarathy and Bartlett, 2012). In contrast to the vertical montage, the horizontal montage is recorded as the difference between bilateral earlobes or mastoids, and is thought to reflect activity from the most peripheral parts of auditory system, including the cochlea and auditory nerve (Galbraith, 1994; Galbraith et al., 2000, 2001; King et al., 2016).

In recent years, researchers have used the widespread availability of whole-scalp multi-channel (e.g., 32-, 64-, or 128-channel) EEG systems either to attempt to identify the different neural generators contributing to the ASSR (e.g., Bidelman, 2015; Tichko and Skoe, 2017; Zhang and Gong, 2017), or to simply integrate the information from multiple electrodes to produce a more accurate or efficient estimate of the overall ASSR (e.g., Bharadwaj and Shinn-Cunningham, 2014; Biesmans et al., 2015; Lu et al., 2020; Luke and Wouters, 2017). The potential benefits of multi-channel EEG for ASSRs in clinical application have also been explored (e.g., van Dun et al., 2007). The availability of data from multiple electrodes or channels raises the question of how best to integrate the information and how to relate multi-channel ASSR results to those from the traditional vertical-montage ASSR. The amplitude of the ASSRs measured through multi-channel and vertical montages may vary substantially (Coffey et al., 2019; Krizman and Kraus, 2019), but the quantitative relationship between them remains unclear.

Regardless of the montage selected, the raw EEG data need to be re-referenced before analysis, and the measured ASSR strength has been found to be affected by the choice of reference. For instance, the spectral magnitude of the ASSR obtained through the traditional vertical montage (e.g., Fz–M1) has been found to be three to four times larger than its magnitude at the same vertex channel when obtained via a multi-channel recording with a common average reference (Fz–CAR) (Bidelman, 2015). Although it was hypothesized that the discrepancy was caused by the contributions from peripheral components on the mastoid channels, i.e., ASSR activation from the cochlea and auditory nerve picked up by the M1 channel (Bidelman, 2015), this explanation remains untested and it is unclear whether the relationship generalizes across a range of ASSR frequencies, or whether it is specific to the region around 100 Hz tested by Bidelman (2015). In addition to affecting the amplitude, the referencing strategy can also affect the ASSR phase coherence between a pair of electrodes (Essl and Rappelsberger, 1998; Fein et al., 1988; Guevara et al., 2005; Hu et al., 2010), as well as the phase estimate at a single electrode (Thatcher, 2012).

Finally, there are choices to be made regarding how to quantify the ASSR strength. The vertical-montage EEG can be quantified through the magnitude and the phase of the periodic ASSR signal. One measure is the spectral magnitude at the ASSR frequency from the Fourier-transformed time waveform, after averaging the time waveform across trials (Batra et al., 1986; Greenberg et al., 1987). Another common measure of ASSR strength is the phase-locking value (PLV, or inter-trial phase coherence) (Stapells et al., 1987). The PLV is obtained by taking the Fourier transform of the waveform from each trial, extracting the unit vector and then calculating the complex average across trials of the unit vector to extract its magnitude. A PLV of 0 corresponds to random noise and 1 corresponds to perfect phase locking, with the phase remaining consistent across all trials (Tallon-Baudry et al., 1996). When compared with spectral magnitude, the PLV has been found to be a more reliable measure (Zhu et al., 2013), and has been widely used in recent ASSR research (Coffey et al., 2016a, 2016b; Mao et al., 2018; Samuelsson et al., 2019; Skosnik et al., 2007; Spencer et al., 2008).

Quantifying the ASSR strength from multi-channel recordings is based on the assumed or estimated phase relationships between channels. Some multi-channel techniques rely on the assumption that the phase of the ASSR remains either equal or opposite across electrodes (Biesmans et al., 2015; van Dun et al., 2009), as would be expected with a single dominant neural generator and a near-instantaneous change in potential at the scalp. This assumption may not always hold, as the scalp ASSR likely reflects a mixture of multiple neural generators with different phases (Bidelman, 2015, 2018; Chandrasekaran and Kraus, 2010; Coffey et al., 2019). To accommodate such phase inconsistencies, more sophisticated measures have been developed, such as the multiple magnitude-squared coherence (MMSC) (da Silva Eloi et al., 2018), complex principal component analysis (cPCA) (Bharadwaj and Shinn-Cunningham, 2014; Lu et al., 2020), and the Hotelling's T^2 (Picton et al., 1987). Finally, an alternative approach to accommodate the possibility of unequal phases is to independently compute ASSR strength within each single channel and to simply summarize the output from all electrodes by averaging or computing other statistics such as root-mean-square (RMS) value. The signal-to-noise ratio (SNR) of the RMS of PLV across all channels has generally been found to be significantly higher than the PLV based on a single channel (Bharadwaj and Shinn-Cunningham, 2014; Lu et al., 2020).

In summary, the recording and analysis of the ASSR via EEG include multiple key factors that may affect the final estimated ASSR strength: stimuli modulation rate, EEG electrode number and placement, choice of reference, and quantification technique. No systematic evaluation of all these factors and their potential combinations has yet been undertaken to evaluate the optimal setup and analysis procedures for ASSR studies. This lack of any systematic study of how best to evaluate multi-channel EEG data has led to difficulties in quantitatively comparing results from different approaches, and uncertainty regarding how to optimally combine information from multiple channels under different recording conditions. The aim of the current study was to provide a systematic approach to tackling these questions. We analyzed an existing EEG dataset that recorded ASSR evoked by modulation rates of around 40, 100, and 200 Hz, to capture the range from those thought to be dominated by cortical responses (40 Hz) to those dominated by subcortical responses (200 Hz). To study the effect of different pre-processing techniques, we compared the phase and strength of ASSR from a vertical-montage approach with various reference strategies. We anticipated that the pre-processing technique used would affect the phase and the strength of the ASSR, estimated through spectral magnitude and PLV. Using this dataset, we also evaluated the differences in response strength and reliability when using single-channel versus multi-channel methods. However, a major difficulty to compare the

output from various methods is their inconsistent scales. For example, the output from one method could be much higher than the upper limit of the other method. Therefore, we designed a standardization process to compare the method output at the AM frequency (signal) with those at neighboring frequencies without any AM (noise floor). Our hypothesis was that multi-channel methods should need fewer trials to distinguish the ASSR at the AM frequency from the noise floor than single-channel methods. To provide a comprehensive evaluation of different ASSR measurement methods, we compared the standardized results from both single- and multi-channel approaches using different referencing strategies and compared the strength of the ASSR with different numbers of trials, ranging from 200 to 1000.

Overall, the aim of this study was to systematically evaluate the effect of various analysis decisions on the strength of the ASSR for single and multi-channel data. We compared two reference strategies on single- and multi-channel data with a variety of EEG analysis methods.

2. Materials and methods

2.1. Data

The data used for these analyses were collected as part of the protocol described in a recently published study (Mehta et al., 2021). The current dataset included recordings from fourteen young normal-hearing participants (4 males and 10 females; age range 18–34 years) who had all provided written informed consent and completed protocols approved by the Institutional Review Board of the University of Minnesota. All participants had audiometric thresholds less than 20 dB HL at octave frequencies between 250 and 8000 Hz. The monaural stimuli for evoking the ASSR included five simultaneously presented tones, all spaced 5/11 octaves apart from their nearest neighbors, and all with a total duration of 437.52 ms, including 10-ms raised-cosine ramps; see Fig. 1. In session one, the central tone was centered around 1806 Hz (with a ±1/4 octave frequency rove between trials) and was sinusoidally amplitude modulated with the sum of two modulators at 43.43 Hz and 98.28 Hz, each with 50% modulation depth. A depth of 50% was used to avoid overmodulation when the two modulators were added, so that their sum never exceed 100% depth. The four flanking tones (two above and two below the central tone) were all coherently sinusoidally amplitude modulated with the sum of 34.28 Hz and 91.42 Hz, each with 25% modulation depth. A lower depth was selected to make the ASSR strength of the single target more similar to the ASSR strength of the four flankers. In session two, the central tone was centered around 3612 Hz (with a ±1/4 octave frequency rove between trials) and was sinusoidally amplitude modulated at 43.43 Hz and 233.13 Hz, each with 50%
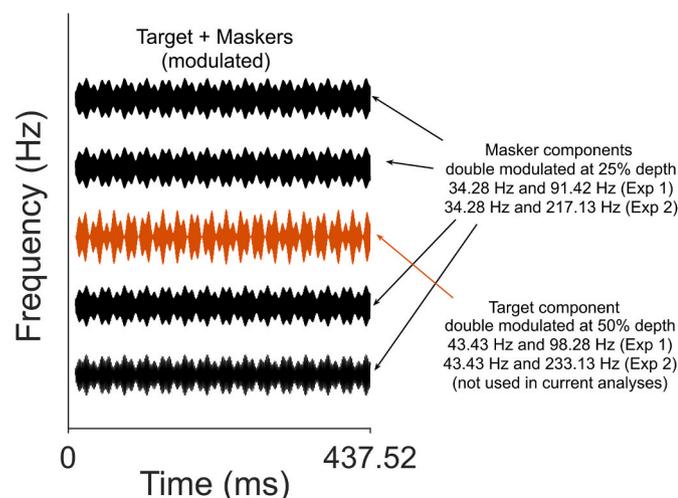
modulation depth, and the four flanking tones were coherently sinusoidally amplitude modulated at 34.28 Hz and 217.13 Hz, each with 25% modulation depth.

In addition to its ready availability, this dataset was chosen as being representative of many that involve compromises between the strong ASSR produced by full modulation and long-duration stimuli (Picton et al., 2003) and the weaker ASSR produced by stimuli that can be used in perceptual experiments while measuring responses from multiple sources simultaneously. It is in such cases that the efficient extraction of the ASSR is of particular importance. In the original study (Mehta et al., 2021), EEG responses to the AM stimuli were recorded with and without an unmodulated precursor stimulus to investigate neural correlates of auditory enhancement. Here, only the trials without a precursor were analyzed. Each condition included 1000 trials. Because the ASSR evoked by the one central tone was not always detectable at high AM rates, only the ASSRs evoked by the four flanking tones with 25% modulation depth were analyzed here. Thus, the ASSR modulation rates of interest were at 34.28 Hz (Sessions 1 and 2), 91.42 Hz (Session 1), and 217.13 Hz (Session 2). Because the 34.28-Hz rate was present in both sessions, a total of 2000 trials were obtained with that frequency. No significant differences in the 34.28-Hz ASSR were detected between sessions 1 and 2 for either reference choice (LMR or CAR) and across multiple recording electrode (Fz, Cz, M1, and M2). Therefore, all analyses at 34.28-Hz were conducted on the 1000 trials in session 1 to keep the number of trials consistent for all frequencies. The figures of all analyses generated using all 2000 trials are attached in the supplementary materials (Figs. S1–S5) in comparison with main Figs. 2–6.

2.2. Pre-processing

The EEG pre-processing and averaging was done using the EEGLAB toolbox (Delorme and Makeig, 2004). The raw waveforms were re-referenced in one of two ways: 1) to the average of two mastoid electrodes (linked-mastoid-referenced strategy, LMR), and 2) to the average of all 64 scalp electrodes (common-average-referenced strategy, CAR). The linked mastoid was chosen to reduce potential noise from a single mastoid. In the analyses involving the phase at mastoids, the mastoids channels were re-referenced to CAR, i.e., $(M1+M2)/2 - CAR$. The main reason is that the phase at mastoids can only be analyzed when referenced to CAR, because LMR referenced to itself will be constantly 0, i.e., $(M1+M2)/2 - (M1+M2)/2 = 0$. For each condition, the continuous EEG time series was divided into epochs. The epoch extended from 100 ms before stimulus onset to 437 ms post stimulus onset. The epoched signal was then baseline corrected relative to the 100-ms pre-stimulus baseline and only the 437-ms recording (while the AM stimuli were playing) was used in following analyses. Further analyses were performed in a Python environment with the MNE-python package (Gramfort, 2013).

2.3. Measurement techniques

2.3.1. Spectral magnitude and phase comparisons of vertical-montage and multi-channel ASSR

The spectral magnitude, phase coherence, and delay were estimated under the two different types of reference (LMR and CAR). The spectral magnitude was calculated by applying a discrete Fourier transform (DFT) to the averaged EEG time waveform at Cz. Phase coherence is represented by the PLV, which was calculated at each electrode by applying a DFT to the time-domain waveforms from individual trials, extracting the phase at each frequency, and then taking the complex average of unit vectors with the phases from each trial. The 1000 trials from session 1 of 34.28-Hz ASSR, 1000 trials of 91.42-Hz ASSR, and 1000 trials of 217.13-Hz ASSR were used in calculating the PLVs from each of the 64 electrodes across the scalp. To test the effect of sample size, a series of subsamples with sizes ranging from 200 to 1000 trials were drawn without replacement from the total of 1000 trials with the

Fig. 1. Schematic representation of the AM stimuli used for measuring ASSR.
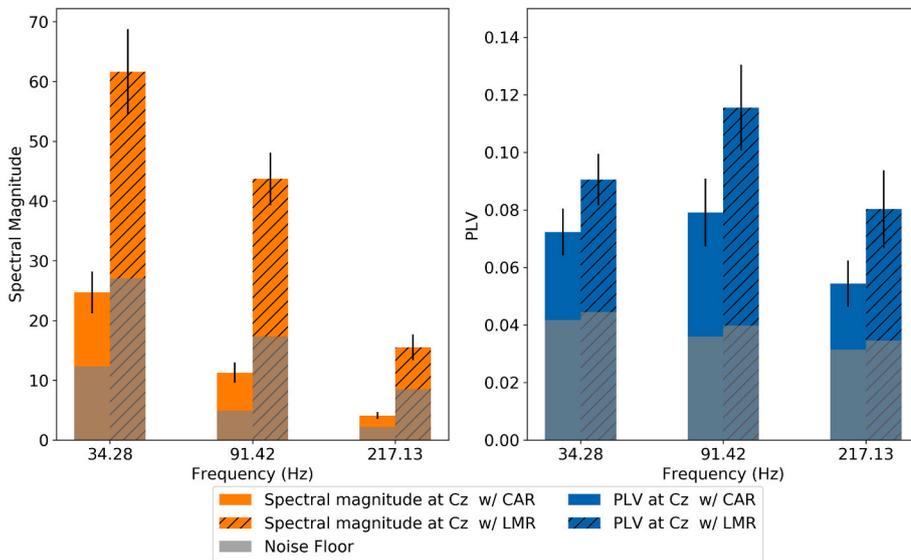
Fig. 2. The local (single-channel at Cz) spectral magnitudes (left panel) and PLVs (right panel) computed with different reference strategies, averaged across participants. The solid bars are calculated with the EEG re-referenced to CAR and the striped bars are calculated with the EEG re-referenced to LMR. The striped bars therefore represent the classical vertical montage. Error bars represent the standard error of the mean across participants. The grey bars represent the noise floor, which is the average spectral magnitude or PLV at unmodulated frequencies surrounding the tagged frequency, as defined in methods section.
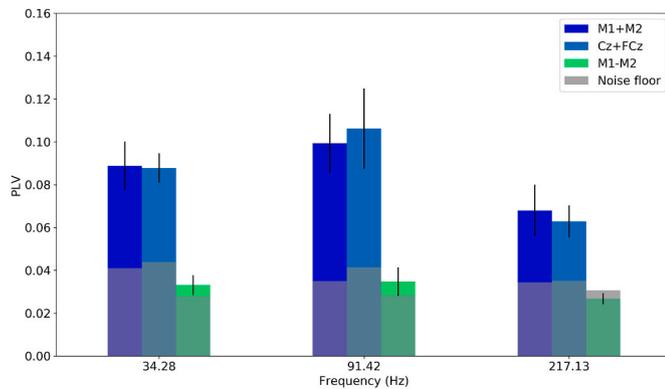
Fig. 3. The group-average PLVs at different locations, re-referenced with the CAR. The bars are clustered by AM frequency. Within a frequency cluster, the dark blue bars show the PLVs measured at mastoids channels (M1+M2); the lighter blue bars show the PLVs measured at vertex channels (Cz + FCz); and the green bars show the PLV of the horizontal difference ASSR (M1-M2). Error bars represent the standard error of the mean across participants. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 4. Circular average of the phase delay between the ASSR phase recorded at the mastoid (M1+M2) and vertex (Cz + FCz) channels with CAR. The horizontal dashed line represents the phase difference equaling to π, i.e., the ASSR recorded at mastoid and vertical channels were in opposite phase. The error bars represent the circular standard deviation of phase differences.

34.28-Hz ASSR and 1000 trials with the 91.42-Hz and 217.13-Hz ASSRs.

2.3.1.1. Quantifying the ASSR. The PLV directly reflects the phase of ASSR and its variance, allowing us to observe the effect of the choice of reference on the single- and multi-channel PLV. Several analysis methods based on different assumptions of the signal and its source(s) have been developed to measure the PLV. Although it is beyond the scope of our study to provide an in-depth examination of the details of the individual multi-channel methods, we focus on evaluating the performance of a standardized version of each of these methods on our dataset.

As mentioned earlier, there are significant differences in the magnitude of the response across these evaluation methods, especially when comparing single-versus multi-channel data. The variety of candidate methods made it challenging to select specific degree of freedom for an F-test. To overcome this issue, we use a z-transformation method that allows for easier comparisons across methods. For any valid analysis method, the estimated ASSR at the tagged AM frequency should be significantly higher than those at untagged frequencies, and an

efficient method can detect this difference with fewer trials. Based on this assumption, the measured ASSR strength was z-transformed by subtracting the mean and dividing by the standard deviation of the ASSR noise floor calculated through each method ($z = (\text{signal} - \text{mean(noise)})/\text{std(noise)}$). The noise floor was estimated at frequencies surrounding each targeted frequency, as follows: For the 34-Hz ASSR, the range was 20–55 Hz; for the 91-Hz ASSR, the range was 75–110 Hz; for the 217-Hz ASSR, the range was 205–240 Hz. In all cases, the stimulus AM frequencies themselves were excluded from the noise-floor calculations.

To better understand the relative advantages and disadvantages of different approaches to deriving the ASSR, the results using the following six different methods of calculating the PLV were compared:

1) Single-channel PLV: The PLV was derived by considering only the signal from a single electrode, Cz.
2) Root mean square of all-channel PLV (PLV$_{\text{RMS}}$): The PLV was derived by calculating the root mean square (RMS) of the PLVs from all electrodes.
3) Time-domain principal component analysis (tPCA): The PLV was derived via time-domain principal component analysis (PCA) by
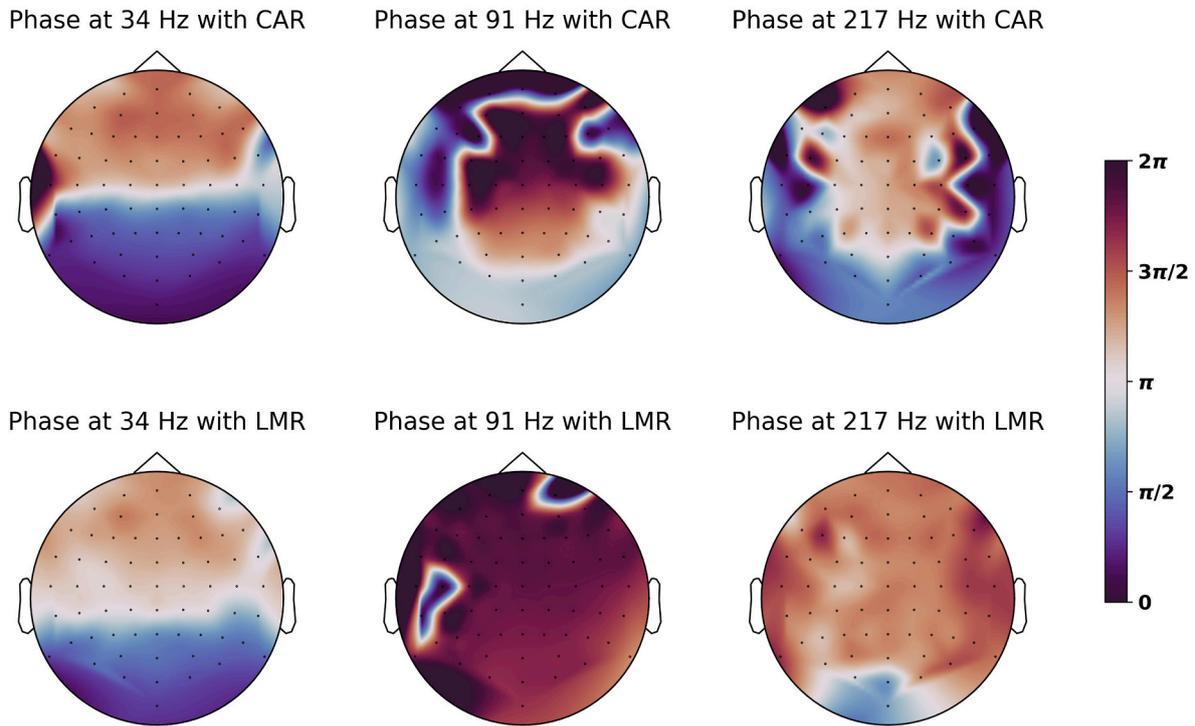
Fig. 5. The topomap of ASSR phase with LMR and CAR. The color represents the circular mean of estimated phase of all participants. With CAR, the values were mostly clustered around one of two opposite phases, as the voltage pattern of dipole under a neutral reference. With LMR, the phases were approximately uniform except for the 34-Hz ASSR. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
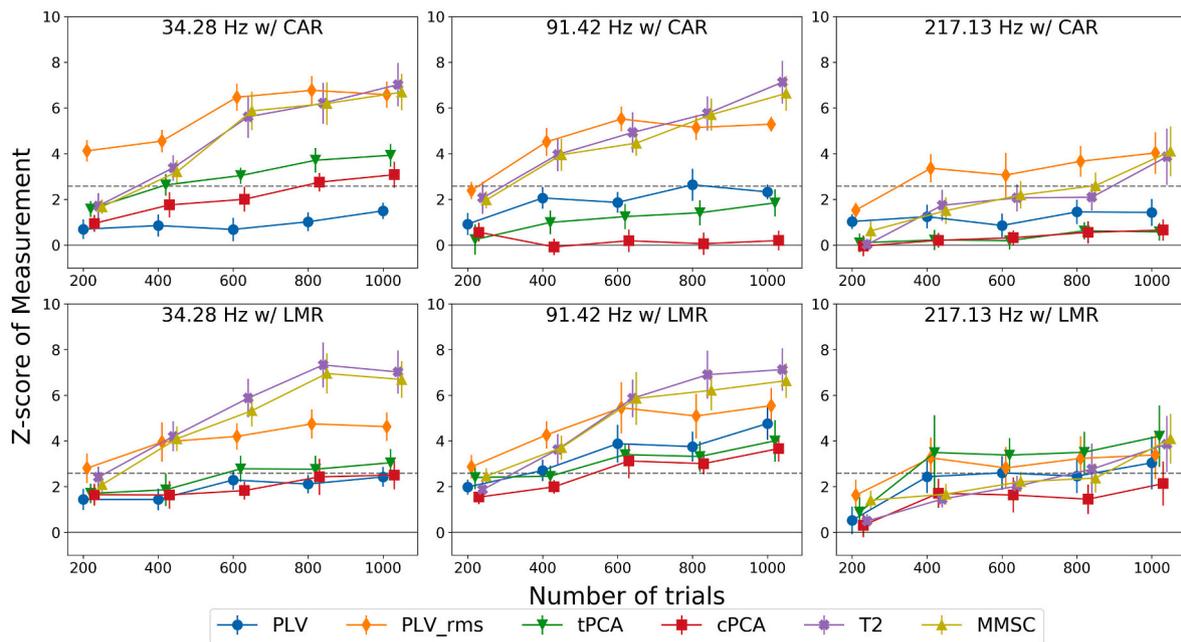
Fig. 6. Average z-scores of the ASSR measured with the various methods. The top row represents the ASSR calculated on the EEG re-referenced to CAR, and the bottom row represents the ASSR calculated on the EEG re-referenced to LMR. The three columns of plots represent the ASSR at 34 Hz, 91 Hz, and 217 Hz, from left to right. The horizontal solid lines represent the z-score where the ASSR is the same as the mean of noise floor. The horizontal dashed lines represent the z-score with 99% confidence rejecting the null hypothesis that there was no ASSR above noise floor. Error bars represent the standard error of the mean z-score across participants.

applying the DFT to the principal component series (i.e., a weighted sum of all 64 channels with real weights obtained through the eigendecomposition of the real temporal covariance matrix) of each trial that explained the most variance.

4) Complex principal component analysis (cPCA): The PLV was derived via spectral-domain complex PCA, using a weighted sum of all 64 channels with the complex weights obtained through the eigendecomposition of the complex cross-spectrum matrix.

5) Hotelling's T^2: The multi-channel Hotelling's T^2 measurement was derived by treating the real and imaginary part of Fourier-transformed 64 channels separately as 128 variables to estimate the 128-by-128 covariance matrix and normalizing the sum of average of 128 variables with the inverted covariance matrix.

6) Multi-channel magnitude squared coherence (MMSC): The multi-channel magnitude squared coherence was derived by normalizing the sum of the average of the Fourier-transformed 64 channels by their inverted complex cross-spectrum matrix.

When referenced to LMR, the first method is equivalent to the classical vertical-montage method, where the ASSR is recorded as the difference between vertex and mastoid electrodes (e.g., Clinard et al., 2010; Hoormann et al., 1992). Methods 2–6 incorporate multi-channel analysis to assess the potential benefits of utilizing high-density EEG over the traditional vertical montage (Method 1). The all-channel PLV_{RMS} method allows inconsistent phases between channels, as the PLVs are computed independently for each channel (Bharadwaj et al., 2015; Mehta et al., 2021; Paul et al., 2017). The tPCA method assumes that the ASSR phases are consistent across channels (either exactly in phase or out of phase), and so the results would be negatively affected by any phase differences between channels except those with exact opposite phases. Although rarely used in multi-channel ASSR analysis, the tPCA method has been used as a baseline condition of multi-channel methods (Bharadwaj and Shinn-Cunningham, 2014). The cPCA method was designed to compensate for the relative phase difference across channels (Bharadwaj and Shinn-Cunningham, 2014), but the cPCA tool developed by Bharadwaj and Shinn-Cunningham (2014) and used in subsequent studies (Paul et al., 2017; Samuelsson et al., 2019; Varghese et al., 2015) actually implemented the all-channel PLV_{RMS} method, rather than a cPCA approach (Lu et al., 2020). Therefore, the performance of an actual cPCA algorithm has yet to be formally tested. In this paper, we developed code that strictly follows the original algorithm of cPCA, where the principal component series were computed as complex weighted sums of Fourier-transformed channels. Both the multi-channel Hotelling's T^2 (e.g. Mijares et al., 2013; Vanheusden et al., 2019) and MMSC (da Silva Eloi et al., 2018; Felix et al., 2018) methods have been used in recent multi-channel EEG studies of ASSR, and both aim to provide an overall estimate of ASSR strength with phase differences between channels corrected by the estimated covariance matrix.

In summary, the six different methods for estimating the ASSR were chosen to represent four typical strategies for treating possible phase inconsistencies across channels: considering only one channel (method 1); independently computing phase coherence for each channel and taking the RMS (method 2); assuming equal or opposite phase across channels (method 3); and estimating and compensating for the phase differences (methods 4–6). The statistical tests were performed in an R environment (Team R Core, 2018) with nlme (Pinheiro et al., 2019), lsmeans (Lenth, 2016), and circular (Agostinelli and Lund, 2017) packages. The code for performing all analyses and plotting figures is available on GitHub (https://github.com/HaoLu-a/MultiPhase_Code).

3. Results

3.1. Spectral magnitude and phase coherence under LMR and CAR

The spectral magnitudes and PLVs at Cz were calculated with the EEG referenced to either CAR or LMR. As noted earlier, the single-channel spectral magnitude and PLV with LMR is equivalent to the classic vertical-montage recording method. The results are shown in Fig. 2. For all combinations of stimulus frequency and reference strategy, both the spectral magnitude and the PLV were significantly above noise floor (p < 0.003 and p < 0.018, respectively, for all conditions).

To explore the effect of AM frequency and reference strategy on the spectral magnitude, a repeated-measures ANOVA was performed on the spectral magnitudes with factors of reference strategy (CAR and LMR) and AM frequency (34, 91, and 217 Hz). Results showed significant main effects of reference strategy ($F_{1,65} = 82.4$, p < 0.001) and AM frequency ($F_{2,65} = 42.3$, p < 0.001), and their interaction ($F_{2,65} = 7.05$, p = 0.002). The spectral magnitude with LMR was significantly higher than that with CAR when the AM frequency was 34 Hz ($t_{65} = 7.19$, p < 0.0001), 91 Hz ($t_{65} = 6.32$, p < 0.0001), or 217 Hz ($t_{65} = 2.22$, p = 0.03). The ratio between the means of spectral magnitude with LMR and CAR was 2.73 at 34 Hz, 3.88 at 91 Hz, and 3.78 at 217 Hz, which is similar to the factor of 3–4 reported by Bidelman (2015). When using CAR, the spectral magnitude at 34 Hz was significantly higher than that at 91 Hz ($t_{65} = 4.01$, p = 0.005, corrected with Tukey's method) and 214 Hz ($t_{65} = 4.01$, p = 0.005, corrected with Tukey's method). When using LMR, the spectral magnitude decreased as the frequency increased, and the differences between all pairs were significant (p < 0.003 for all, corrected with Tukey's method).

To explore the effect of AM frequency and reference strategy on the measured PLV, a repeated-measures ANOVA was performed on the PLVs with factors of reference strategy (CAR and LMR) and AM frequency (34, 91, and 217 Hz). Results showed a significant main effect of reference strategy ($F_{1,65} = 9.96$, p = 0.002) and AM frequency ($F_{2,65} = 4.14$, p = 0.02), but their interaction was not significant ($F_{2,65} = 0.39$, p = 0.68). The PLVs were significantly larger when re-referenced to LMR than to CAR ($t_{65} = 3.16$, p = 0.002). The 91-Hz PLV at Cz was significantly larger than the 217-Hz ($t_{65} = 2.88$, p = 0.015, p values were adjusted by Tukey's method) PLV. There was no significant difference between the other pairs (p values were adjusted by Tukey's method). The consistent decrease in spectral magnitude (but not PLV) with increasing modulation frequency is consistent with the expectation that the overall EEG response (but not necessarily its phase consistency) decreases in roughly inverse proportion to frequency (Buzsáki, 2006). Having established the expected pattern of responses with spectral magnitude (and its dependence on the re-reference strategy), the remainder of our study focuses on the PLV.

To compare the strength of the ASSR at various locations on the scalp, PLVs were calculated with CAR using three different montages of LMR (M1+M2), vertex electrodes (Cz + FCz), and the horizontal difference (M1-M2). A previous study explained the difference in ASSR strength measured with CAR and LMR in terms of the mastoid channels in the LMR conditions picking up signals from the cochlear microphonic and auditory nerve (Bidelman, 2015). If the hypothesized contamination from peripheral components exists, such signals should be detected via the horizontal difference (M1-M2) (Billings et al., 2019; Galbraith et al., 2000; King et al., 2016). The results are shown in Fig. 3. The PLVs were first compared with the PLV noise floor, as defined in methods section. Paired t-tests showed that the PLVs at all three AM frequencies were significantly higher than the noise floor when recorded as M1+M2 (p < 0.012 for all three frequencies) and Cz + FCz (p < 0.002 for all three frequencies). In contrast, the PLVs measured as M1-M2 were not significantly different from the noise floor (p > 0.12 for all three frequencies), suggesting that no significant horizontally oriented components were recorded at the mastoids. Thus, the contribution of peripheral components seems unlikely to explain why the ASSR amplitude is larger for the LMR than for the CAR referencing strategy. A repeated-measures ANOVA was performed on the PLVs with electrode montage (M1+M2, Cz + FCz, or M1–M2) and frequency (34, 91, or 217 Hz) as factors. Results showed significant main effects of electrode montage ($F_{2,104} = 27.7$, p < 0.0001) and frequency ($F_{2,104} = 5.52$, p = 0.0052), but no significant interaction ($F_{4,104} = 0.775$, p = 0.54). Pairwise comparisons showed that the ASSR at 91 Hz was significantly larger than the 217-Hz ASSR ($t_{104} = 3.29$, p = 0.0039), and the M1–M2 horizontal ASSR was significantly smaller than the M1+M2 and Cz + FCz ASSRs (p < 0.0001 for both). Other pairwise comparisons showed no significant differences between pairs of frequencies or locations (p > 0.10 for all).

3.2. Phase delay across electrodes at different locations

Given that peripheral components seem unlikely to explain the differences in ASSR due to referencing strategy (CAR and LMR), the phase differences between the ASSR recorded at M1+M2 and Cz + FCz were also examined and are shown in Fig. 4. The results of a Rayleigh test of uniformity were used to reject the null hypothesis of uniform distribution between 0 and 2π rad for the phase of the ASSR at 34 Hz (z = 0.61, p = 0.0041), 91 Hz (z = 0.697, p < 0.0001), and 217 Hz (z = 0.472, p = 0.041). Therefore, the circular mean of phase difference for all three AM frequencies should be a meaningful estimate of the actual phase delay between mastoid and vertex, implying that the ASSRs recorded at vertex (Cz + FCz) and mastoid (M1+M2) channels were approximately in opposite phase (π rad) for all three tested frequencies.

To investigate the effect of re-referencing strategy on the phases of ASSR on the scalp, the phases at all electrodes locations were estimated in spectral domain and were averaged across participants. The results are plotted as topomaps in Fig. 5. The phase patterns with CAR for three AM frequencies reflected the different contributions from cortical and subcortical sources. For 34-Hz ASSR, the patterns of the phase topomaps were similar when using CAR and LMR, where the phases in the frontal and occipital regions were approximately opposite to each other. For 91-Hz and 217-Hz ASSR, the phases of parietal and temporal electrodes were approximately opposite to each other when re-referenced to CAR but were more nearly uniform when re-referenced to LMR.

3.3. Comparison of performance of measurements in detecting ASSR

To compare the efficiency of the various methods in detecting the presence of an ASSR, the outputs were z-transformed to unify their scales, as described in the methods section. The raw outputs from various methods before z-transformation are plotted in Fig. S6 and Fig. S7. The resulting z-scores are plotted in Fig. 6. To control the total number of factors within each model and ensure the interpretability of interaction effects, repeated-measures ANOVAs were applied to the z-scores at the three frequencies separately, with ASSR method, sample size, and reference strategy as independent variables. As the relationship between the average z-score and sample size was approximately linear, the sample size was treated as a continuous variable to reduce the number of parameters to be estimated.

For the z-scores of the 34-Hz ASSR, significant main effects of method ($F_{5,803}$ = 4.88, p < 0.0001) and sample size ($F_{1,803}$ = 207, p < 0.0001) were found, along with a significant interaction between method and sample size ($F_{5,803}$ = 14.8, p < 0.0001). Other effects and interactions were not significant. As there was a significant interaction between the continuous variable of sample size and the discrete variable of method, pairwise comparisons were separately performed on the intercept and slope of the linear relationship between z-score and sample size with the various methods. The intercept was the z-score estimated at 600 trials for each condition to compare the overall performance of the methods, and the slope represented the benefit obtained with a given method by increasing the sample size. The results of pairwise comparisons on the intercept were: {PLV_{rms}, MMSC, T^2} > {$PLV_{single-channel}$, tPCA, cPCA} (p > 0.10 for methods within a set; p < 0.0001 for all pairs between sets; all p values were adjusted with Tukey's method). The results of pairwise comparison on the slope were: {MMSC, T^2} > {$PLV_{single-channel}$, PLV_{rms}, tPCA, cPCA} (p > 0.27 for methods within a set; p < 0.0004 for all pairs between sets; all p values were adjusted with Tukey's method).

For the z-scores of the 91-Hz ASSR, the test showed significant main effects of method ($F_{5,803}$ = 2.97, p = 0.0115) and sample size ($F_{1,803}$ = 184, p < 0.0001), and significant interactions between method and sample size ($F_{5,803}$ = 11.2, p < 0.0001). Other effects were not significant (p > 0.072 for all). Due to the significant interaction between sample size and reference strategy, pairwise comparisons between the intercept and slope of the linear relationship between z-score and

sample sizes with various methods were separately conducted for CAR and LMR. For CAR, the results of pairwise comparison on the intercept were: {MMSC, T^2, PLV_{rms}} > {$PLV_{single-channel}$, tPCA} > cPCA (p > 0.21 for methods within a set; p < 0.001 for all pairs between sets; all p values were adjusted with Tukey's method), except that tPCA > cPCA(p = 0.051). The results of pairwise comparison on the slope were: {MMSC, T^2} > { PLV_{rms}, $PLV_{single-channel}$, tPCA, cPCA} (p > 0.72 for methods within a set; p < 0.009 for all pairs between sets; all p values were adjusted with Tukey's method), except that MMSC > PLV_{rms} (p = 0.086).

For the z-scores of the 217-Hz ASSR, the test showed significant main effects of method ($F_{5,803}$ = 2.71, p = 0.019) and sample size ($F_{1,803}$ = 84.4, p < 0.0001), and their interaction ($F_{5,803}$ = 2.93, p = 0.012). No other effects were significant (p > 0.23 for all). The results of pairwise comparison on the intercept were: PLV_{rms} > {MMSC, T2, tPCA, $PLV_{single-channel}$} > cPCA (p > 0.23 for methods within a set; p < 0.018 for all pairs between sets; all p values were adjusted with Tukey's method), except that PLV_{rms} > MMSC (p = 0.062). The results of pairwise comparison showed that the slope of cPCA was significantly lower than T2 (p = 0.034). The difference between all other pairs was insignificant (p > 0.069 for all).

In summary, most multi-channel methods were able to detect ASSR with fewer trials than the single-channel method. The exceptions were the tPCA and cPCA methods, which were less efficient. The effect of re-referencing strategy was only significant for the 91-Hz ASSR.

4. Discussion

4.1. Referencing strategy affects the strength of ASSR measured by phase coherence

Our results show a decrement in both spectral magnitude and PLV when changing from LMR to CAR (Fig. 2) for AM frequencies ranging from around 40 Hz–200 Hz, Therefore, the strength of ASSR recorded with a single or linked mastoid reference (King et al., 2016; Varghese et al., 2015) should not be directly compared with that of ASSR recorded with relatively neutral reference, such as the CAR (Bidelman, 2015; Herdman et al., 2002; Ross et al., 2005) or the 7th cervical vertebra (Ananthakrishnan et al., 2016; Swaminathan et al., 2008).

Switching from CAR to LMR resulted in an increase in spectral magnitudes at all frequencies by a factor of around three, as also found by Bidelman (2015). Although the PLV also showed an increase when switching from LMR to CAR (Fig. 2), our phase coherence results do not support the previous interpretation that it was caused by the influence of peripheral (cochlear or auditory-nerve) components at the mastoid channels (Bidelman, 2015). In fact, we found the PLVs at mastoid (M1+M2) and vertex (Cz + FCz) under CAR to be similarly large, and the PLV from the horizontal montage of ASSR (M1-M2), which has been used as a measure of the peripheral component of ASSR (Billings et al., 2019; Galbraith, 1994; Galbraith et al., 2000), to be not significantly different from the noise floor. Therefore, the effect of re-referencing strategy on the measured ASSR strength seems unlikely to be due to the contribution from peripheral components.

To provide an alternative explanation of the effect of re-referencing strategy, we examined the pattern of ASSR phase across the scalp (Figs. 4 and 5). The phase difference of the ASSR at vertex and mastoids was approximately half a period (π rad) for all three frequencies (Fig. 4), as has been reported in earlier studies (Herdman et al., 2002; Huis et al., 1977). This pattern may be the result of recording from two sides of the same dipole. The vertical dipole located in inferior colliculus (IC) has been confirmed in various ASSR studies as one of the main sources of ASSR (Bidelman, 2015, 2018; Chandrasekaran and Kraus, 2010; Coffey et al., 2016; Herdman et al., 2002) and aligns with the direction of the vertical-montage electrodes. For this reason, the ASSR recorded with a vertical montage (Fz-M1 or Fz-LMR) is the result of summing the magnitudes of the opposite-phase signals at Fz (Fz-CAR) and M1 (M1-CAR),

because CAR can be considered approximately neutral with adequate electrode coverage (Dien, 1998). Thus, the enhanced amplitude of the ASSR when measured with the vertical montage can be considered a consequence of the spatial sampling of two points that are in opposite phase, compared with sampling just one of the points relative to neutral (Fig. 7). In addition to the frequencies tested in our study, another study has shown that envelope frequencies lower than 20 Hz evoked ASSR at Cz and TP7 in opposite phases, as a result of the vertical source at auditory cortex (Aiken and Picton, 2008). Although this explanation provides a good qualitative account of the data, it should be noted that the phases of ASSR at 34 Hz and 217 Hz were only marginally significantly different from uniform distribution. It seems likely that the ASSR around 40 Hz involves contributions from cortical sources in addition to the vertical IC source (Farahani et al., 2017; Herdman et al., 2002; Poulsen et al., 2009; Spencer, 2012), as indicated by the different phase pattern found on the topomap (Fig. 5). However, we are unable to provide a valid delay analysis to support this hypothesized interpretation as the length of the stimulus was too short to implement a reasonable delay analysis. At around 200 Hz, the lack of strong evidence of opposite phases between vertex and mastoid channels in our data may be due to the decreased signal-to-noise ratio of the vertical ASSR, as expected at high stimulus frequencies (Galbraith et al., 2000). In summary, the ASSR recorded at Cz + FCz and M1+M2 likely reflects contributions from the two sides of the same vertical dipole(s).

4.2. Choosing an appropriate re-referencing strategy for ASSR research

When comparing the LMR and CAR re-referencing strategies, our results confirm that the mastoid channels actively measure the ASSR. However, in most cases, the CAR cannot be considered completely neutral either: most EEG systems only cover the scalp rather than the entire head surface, and the sum of voltage changes of such incomplete coverage area may not be consistently zero, leading to a biased CAR. This effect, termed the polar average reference effect, is not negligible, even for a 270-degree coverage by the EEG system (Junghöfer et al., 1999). Nevertheless, the CAR remains more neutral than the LMR, and if

the research analyses aim to utilize the dipole model to locate the sources of ASSR, the more neutral reference CAR should be used to preserve the original response patterns at the electrodes.

In a more common situation, where the ASSR needs to be detected through multi-channel EEG, there may be trade-offs between the CAR and LMR. The CAR is the average of all available EEG channels, and this global averaging process naturally smooths out local signal (ASSR at different phases) and noise (bad channels). Therefore, the CAR is closer to the neutral reference, with less noise, but also less signal, than the LMR. The LMR is more susceptible to artifacts affecting either of the mastoid channels, but the ASSR signal at LMR will also enhance the ASSR at vertex channels with approximately opposite phase (Herdman et al., 2002). In other words, the CAR emphasizes minimizing noise while the LMR emphasizes maximizing the signal. Re-referencing to the LMR may help to make the phase of the ASSR across the scalp more uniform (see Fig. 5) and may thus affect the performance of multi-channel methods relying on phase constancy (Biesmans et al., 2015; Luke and Wouters, 2017; van Dun et al., 2009). This constant-phase assumption is derived by assuming the ASSR is evoked by single dipole, and can accommodate out-of-phase components (by assigning a negative amplitude to those channels); however, in cases where multiple dipoles contribute significantly to the signal, such a simplification may not be appropriate (Bharadwaj and Shinn-Cunningham, 2014). The following section evaluates the performance of the different multi-channel techniques tested.

4.3. Performance of various multi-channel methods under CAR and LMR

Our results revealed significant differences in performance between the various multi-channel methods we tested. The tPCA method was unable to reliably detect the ASSR at 91 Hz and 217 Hz with all 1000 trials with CAR but was able to detect ASSR at all frequencies with LMR. This improvement is likely the result of the enhanced ASSR signal and, more importantly, the relatively uniform phase across channels that was found after re-referencing to the LMR. Therefore, other multi-channel ASSR methods assuming consistent phase across channels (Biesmans
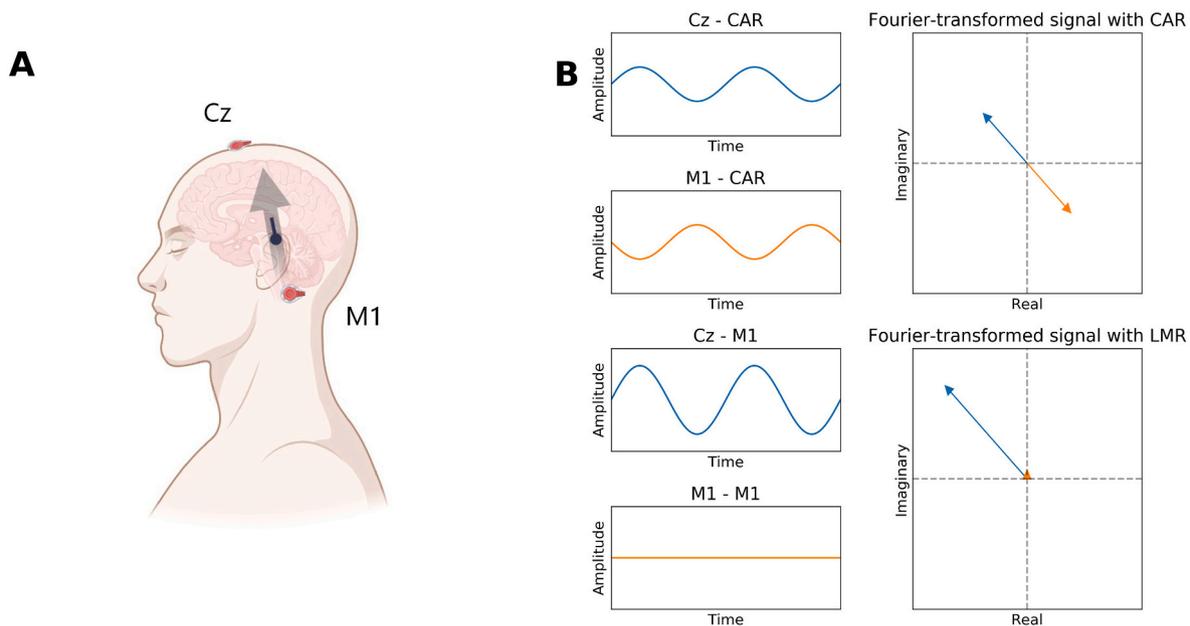
Fig. 7. The hypothesized interpretation of the enhanced ASSR when re-referenced to mastoid channels. Panel A shows the spatial relationship between the vertical-montage active electrodes (Fz & M1) and the brainstem vertical dipole. The brainstem dipole is approximately vertical. Panel B shows the ASSR signal recorded at Fz and M1 with CAR and LMR. The top two rows of panel B are recorded ASSR with CAR, and the bottom two rows are ASSR with LMR. It was assumed that LM was identical to M1. When ASSR was referenced to CAR, the phase of recorded ASSR was opposite as the CAR was approximately neutral. When ASSR was referenced to LMR, there was no signal at mastoid but the strength at Fz was enhanced. From another point of view, Fz – M1 = (Fz – CAR) – (M1 – CAR). Figure created with BioRender.com.

et al., 2015; Luke and Wouters, 2017; van Dun et al., 2009) may also benefit from re-referencing the data to the LMR when the phase of ASSR at mastoids is opposite to that around vertex electrodes.

Although tPCA benefited from LMR, the performance of PCA methods was still generally inferior to that of the other methods. The cPCA method was consistently worse than the other two methods accommodating for phase differences (MMSC and Hotelling's T^2) and showed little or no advantage over the single-channel PLV or even tPCA, when applied by strictly following the theory proposed by Bharadwaj and Shinn-Cunningham (2014). The similar pattern between the z-scores using cPCA and tPCA suggested that cPCA could not provide better performance by compensating for potential phase differences, irrespective of re-referencing method. In fact, the generally poor performance of both the cPCA and tPCA methods reveals a potential problem in applying PCA to analyze the ASSR. The PCA method was designed to decompose a multidimensional input into orthogonal principal components ranked by the variance explained by each component (Wold et al., 1987). Through dropping components that explained little variance of input, the dimensionality of original signal was reduced, while the most variance was kept. As Bharadwaj and Shinn-Cunningham (2014) proposed, our results were computed by keeping only the most dominant single principal component for both cPCA and tPCA, so multi-channel EEG can be summarized as a single estimated value for each frequency. Due to the typically low signal-to-noise ratio of ASSR, there is no guarantee that the dominant principal component will represent the ASSR, so a typical PCA may not be ideal for extracting the ASSR embedded in multi-channel recording. To provide an intuitive insight into the poor performance of tPCA and cPCA, the average ratio of variance explained by the first component was 40.7% for tPCA and 55.6% for cPCA while that for the top 5 components was 76.7% for tPCA and 83.4 for cPCA, suggesting that only selecting the first component was insufficient to represent the multi-channel ASSR data. Further investigation on refining criteria for picking the component which captures the ASSR may improve the performance of cPCA and tPCA but may also add to the difficulty in setting up replicable EEG data analysis protocols. Besides, an important underlying assumption of PCA is that those components are orthogonal, which also requires careful validation for periodic signals like the ASSR. Therefore, based on both empirical and theoretical grounds, the application of PCA to multi-channel EEG data cannot be currently recommended to estimate the strength of ASSR.

In almost all conditions, the relatively simple PLV$_{rms}$ method is among the methods that require the smallest number of trials to reliably detect the ASSR above the noise floor. At 34 Hz and 91 Hz, the z-score of PLV$_{rms}$ reached the threshold for rejecting null hypothesis to detect ASSR with less than 200 trials, which was less than a fifth of the number of trials needed for single-channel PLV, suggesting the potential reduction in necessary recording time compared with traditional single-channel method. Even at 217 Hz, when the ASSR was almost undetectable with all 1000 trials for single-channel method, the z-score of PLV$_{rms}$ reached threshold within 400–600 trials. As shown in Figs. S6 and S7, the PLV$_{rms}$ at AM frequency was similar or even less than the single-channel PLV. The main reason of the advantage of PLV$_{rms}$ is likely due to the reduced variance at the untagged frequencies, i.e., the all-channel RMS of PLV provided smoother noise floor, making it easier to detect ASSR. We predict that this efficiency can be further improved with more electrodes (e.g., 128 or 256 electrode EEG system). A similar advantage in efficiently detecting ASSR with relatively fewer trials can also be achieved by the MMSC and Hotelling's T^2 methods, but both require estimating and inverting the covariance matrix, which can be more computationally expensive than the simple PLV$_{rms}$, and that computational complexity grows cubically with the number of electrodes. Additionally, as the PLV$_{rms}$ method requires relatively fewer trials, it is a simple and sensitive solution to reduce the overall recording time for longer experimental designs with multiple conditions. Therefore, the PLV$_{rms}$ is generally recommended for efficiently detecting ASSR with multi-channel EEG. As the number of trials increases to 1000, the

ASSR signal recorded through MMSC and Hotelling's T^2 is largest compared with the noise floor for the 40- and 100- Hz stimulus. With sufficient data and computational resources, researchers may therefore wish to use methods that allow for unequal phases to obtain maximal signal-to-noise ratio for recorded ASSR at 40 Hz and 100 Hz.

Our evaluation also needs to be interpreted with caution, as it was designed to examine the efficiency through comparing tagged AM frequency and untagged frequencies. To standardize the different scales of various methods, we performed a z-transform, which relies on the relative difference between ASSR at AM frequency and noise floor, and it did not include any comparison between the absolute values. Thus, although the method is most efficient at detecting the presence of the ASSR, it may not be the most effective at determining the strength of a specific source or dipole.

5. Conclusions

When interpreting the ASSR recorded with a vertical montage, the mastoid channels should not be treated as being electrically neutral, especially for high-frequency ASSRs ($\geq$100 Hz), which include a substantial contribution from subcortical sources. As multi-channel ASSRs become more common, it is crucial to take the non-neutral property of mastoid channels into consideration while analyzing EEG data. In experiments aiming to detect reliable ASSR signals, we generally recommend using the CAR and calculating the RMS of the PLV across all channels for its efficiency and robustness. The performance of the PLV$_{RMS}$ in detecting ASSR was generally not affected by the reference strategy, so the CAR is preferred for its robustness against local noise. However, other methods can be considered when prior knowledge concerning the phase relationships between electrodes is available. When the ASSR phase at mastoid is reliably opposite to the ASSR phase at vertex channels, applying LMR will improve the performance over CAR for methods that rely on the assumption that the ASSR phase is constant across all electrodes.

CRediT authorship contribution statement

Hao Lu: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Visualization. **Anahita H. Mehta:** Software, Validation, Investigation, Resources, Data curation, Writing – review & editing, Visualization. **Andrew J. Oxenham:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crneur.2022.100061.

References

Agostinelli, C., Lund, U., 2017. R Package "Circular": Circular Statistics, version 0.4-93. https://r-forge.r-project.org/projects/circular/.

Aiken, S.J., Picton, T.W., 2008. Human cortical responses to the speech envelope. Ear Hear. 29 (2), 139–157. https://doi.org/10.1097/AUD.0b013e31816453dc.

Ananthakrishnan, S., Krishnan, A., Bartlett, E., 2016. Human frequency following response. Ear Hear. 37 (2), e91–e103. https://doi.org/10.1097/AUD.0000000000000247.

Batra, R., Kuwada, S., Maher, V.L., 1986. The frequency-following response to continuous tones in humans. Hear. Res. 21 (2), 167–177. https://doi.org/10.1016/0378-5955(86)90037-7.

Bharadwaj, H.M., Lee, A.K.C., Shinn-Cunningham, B.G., 2014. Measuring auditory selective attention using frequency tagging. Front. Integr. Neurosci. 8 (FEB), 6. https://doi.org/10.3389/fnint.2014.00006.

Bharadwaj, H.M., Masud, S., Mehraei, G., Verhulst, S., Shinn-Cunningham, B.G., 2015. Individual differences reveal correlates of hidden hearing deficits. J. Neurosci. 35 (5), 2161–2172. https://doi.org/10.1523/JNEUROSCI.3915-14.2015.

Bharadwaj, H.M., Shinn-Cunningham, B.G., 2014. Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. Clin. Neurophysiol. 125 (9), 1878–1888. https://doi.org/10.1016/j.clinph.2014.01.011.

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. https://doi.org/10.1016/j.heares.2015.01.011.

Bidelman, G.M., 2018. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. Neuroimage 175, 56–69. https://doi.org/10.1016/j.neuroimage.2018.03.060.

Biesmans, W., Bertrand, A., Wouters, J., Moonen, M., 2015. Optimal spatial filtering for auditory steady-state response detection using high-density EEG, 2015-Augus. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp. 857–861. https://doi.org/10.1109/ICASSP.2015.7178091.

Billings, C.J., Bologna, W.J., Muralimanohar, R.K., Madsen, B.M., Molis, M.R., 2019. Frequency following responses to tone glides: effects of frequency extent, direction, and electrode montage. Hear. Res. 375, 25–33. https://doi.org/10.1016/j.heares.2019.01.012.

Buzsáki, G., 2006. Rhythms of the Brain. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301069.001.0001.

Chandrasekaran, B., Kraus, N., 2010. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47 (2), 236–246. https://doi.org/10.1111/j.1469-8986.2009.00928.x.

Clinard, C.G., Tremblay, K.L., Krishnan, A.R., 2010. Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. Hear. Res. 264 (1–2), 48–55. https://doi.org/10.1016/j.heares.2009.11.010.

Coffey, E.B.J., Colagrosso, E.M.G., Lehmann, A., Schönwiesner, M., Zatorre, R.J., 2016a. Individual differences in the frequency-following response: relation to pitch perception. PLoS One 11 (3), e0152374. https://doi.org/10.1371/journal.pone.0152374.

Coffey, E.B.J., Herholz, S.C., Chepesiuk, A.M.P., Baillet, S., Zatorre, R.J., 2016b. Cortical contributions to the auditory frequency-following response revealed by MEG. Nat. Commun. 7 (1), 11070 https://doi.org/10.1038/ncomms11070.

Coffey, E.B.J., Nicol, T., White-Schwoch, T., Chandrasekaran, B., Krizman, J., Skoe, E., Zatorre, R.J., Kraus, N., 2019. Evolving perspectives on the sources of the frequency-following response. Nat. Commun. 10 (1), 5036. https://doi.org/10.1038/s41467-019-13003-w.

Cone-Wesson, B., Dowell, R.C., Tomlin, D., Rance, G., Ming, W.J., 2002. The auditory steady-state response: comparisons with the auditory brainstem response. J. Am. Acad. Audiol. 13 (4), 173–187.

da Silva Eloi, B.F., Antunes, F., Felix, L.B., 2018. Improving the detection of auditory steady-state responses near 80 Hz using multiple magnitude-squared coherence and multichannel electroencephalogram. Biomed. Signal Process Control 42, 158–161. https://doi.org/10.1016/j.bspc.2018.01.017.

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 (1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Dien, J., 1998. Issues in the application of the average reference: review, critiques, and recommendations. Behav. Res. Methods Instrum. Comput. 30 (1), 34–43. https://doi.org/10.3758/BF03209414.

Duarte, J.L., Alvarenga, K. de F., Garcia, T.M., Filho, O.A.C., Lins, O.G., 2008. Auditory steady-state response in the auditory evaluation: clinical application. Pro-Fono 20 (2), 105–110. https://doi.org/10.1590/S0104-56872008000200006.

Essl, M., Rappelsberger, P., 1998. EEG cohererence and reference signals: experimental results and mathematical explanations. Med. Biol. Eng. Comput. 36 (4), 399–406. https://doi.org/10.1007/BF02523206.

Farahani, E.D., Goossens, T., Wouters, J., van Wieringen, A., 2017. Spatiotemporal reconstruction of auditory steady-state responses to acoustic amplitude modulations: potential sources beyond the auditory pathway. Neuroimage 148, 240–253. https://doi.org/10.1016/j.neuroimage.2017.01.032.

Fein, G., Raz, J., Brown, F.F., Merrin, E.L., 1988. Common reference coherence data are confounded by power and phase effects. Electroencephalogr. Clin. Neurophysiol. 69 (6), 581–584. https://doi.org/10.1016/0013-4694(88)90171-X.

Felix, L.B., Antunes, F., da S Carvalho, J.A., Barroso, M.F., dos, S., Miranda de Sá, A.M.F. L., 2018. Comparison of univariate and multivariate magnitude-squared coherences in the detection of human 40-Hz auditory steady-state evoked responses. Biomed. Signal Process Control 40, 234–239. https://doi.org/10.1016/j.bspc.2017.09.024.

Galbraith, G.C., 1994. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. Electroencephalogr. Clin. Neurophysiology Evoked Potentials Sect. 92 (4), 321–330. https://doi.org/10.1016/0168-5597(94)90100-7.

Galbraith, G.C., Bagasan, B., Sulahian, J., 2001. Brainstem frequency-following response recorded from one vertical and three horizontal electrode derivations. Percept. Mot. Skills 92 (1), 99–106. https://doi.org/10.2466/pms.2001.92.1.99.

Galbraith, G.C., Threadgill, M.R., Hemsley, J., Salour, K., Songdej, N., Ton, J., Cheung, L., 2000. Putative measure of peripheral and brainstem frequency-following in humans. Neurosci. Lett. 292 (2), 123–127. https://doi.org/10.1016/S0304-3940(00)01436-1.

Gander, P.E., Bosnyak, D.J., Roberts, L.E., 2010. Evidence for modality-specific but not frequency-specific modulation of human primary auditory cortex by attention. Hear. Res. 268 (1–2), 213–226. https://doi.org/10.1016/j.heares.2010.06.003.

Gerken, G.M., Moushegian, G., Stillman, R.D., Rupert, A.L., 1975. Human frequency-following responses to monaural and binaural stimuli. Electroencephalogr. Clin. Neurophysiol. 38 (4), 379–386. https://doi.org/10.1016/0013-4694(75)90262-X.

Glaser, E.M., Suter, C.M., Dasheiff, R., Goldberg, A., 1976. The human frequency-following response: its behavior during continuous tone and tone burst stimulation. Electroencephalogr. Clin. Neurophysiol. 40 (1), 25–32. https://doi.org/10.1016/0013-4694(76)90176-0.

Gnanateja, G.N., Rupp, K., Llanos, F., Remick, M., Pernia, M., Sadagopan, S., Teichert, T., Abel, T.J., Chandrasekaran, B., 2021. Frequency-following responses to speech sounds are highly conserved across species and contain cortical contributions. Eneuro 8 (6). https://doi.org/10.1523/ENEURO.0451-21.2021. ENEURO.0451-21.2021.

Gorina-Careta, N., Kurkela, J.L.O., Hämäläinen, J., Astikainen, P., Escera, C., 2021. Neural generators of the frequency-following response elicited to stimuli of low and high frequency: a magnetoencephalographic (MEG) study. Neuroimage 231, 117866. https://doi.org/10.1016/j.neuroimage.2021.117866.

Gramfort, A., 2013. MEG and EEG data analysis with MNE-Python. Front. Neurosci. 7, 267. https://doi.org/10.3389/fnins.2013.00267.

Greenberg, S., Marsh, J.T., Brown, W.S., Smith, J.C., 1987. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hear. Res. 25 (2–3), 91–114. https://doi.org/10.1016/0378-5955(87)90083-9.

Guevara, R., Velazquez, J.L.P., Nenadovic, V., Wennberg, R., Senjanović, G., Dominguez, L.G., 2005. Phase synchronization measurements using electroencephalographic recordings: what can we really say about neuronal synchrony? Neuroinformatics 3 (4), 301–313. https://doi.org/10.1385/NI:3:4:301.

Gutschalk, A., Micheyl, C., Oxenham, A.J., 2008. Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol. 6 (6), 1156–1165. https://doi.org/10.1371/journal.pbio.0060138.

Heo, J., Baek, H.J., Hong, S., Chang, M.H., Lee, J.S., Park, K.S., 2017. Music and natural sounds in an auditory steady-state response based brain–computer interface to increase user acceptance. Comput. Biol. Med. 84, 45–52. https://doi.org/10.1016/j.compbiomed.2017.03.011.

Herdman, A.T., Lins, O., van Roon, P., Stapells, D.R., Scherg, M., Picton, T.W., 2002. Intracerebral sources of human auditory steady-state responses. Brain Topogr. 15 (2), 69–86. https://doi.org/10.1023/A:1021470822922.

Holmes, E., Purcell, D.W., Carlyon, R.P., Gockel, H.E., Johnsrude, I.S., 2018. Attentional modulation of envelope-following responses at lower (93–109 Hz) but not higher (217–233 Hz) modulation rates. J. Assoc. Res. Otolaryngol. 19 (1), 83–97. https://doi.org/10.1007/s10162-017-0641-9.

Hoormann, J., Falkenstein, M., Hohnsbein, J., Blanke, L., 1992. The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. Hear. Res. 59 (2), 179–188. https://doi.org/10.1016/0378-5955(92)90114-3.

Hu, S., Stead, M., Dai, Q., Worrell, G.A., 2010. On the recording reference contribution to EEG correlation, phase synchrony, and coherence. IEEE Trans. Syst. Man Cybern. B Cybern. 40 (5), 1294–1304. https://doi.org/10.1109/TSMCB.2009.2037237.

Huis, F., Osterhammel, P., Terkildsen, K., 1977. Frequency following auditory brain stem responses in man. Scand. Audiol. 6 (1), 27–34. https://doi.org/10.3109/01050397709044995.

Junghöfer, M., Elbert, T., Tucker, D.M., Braun, C., 1999. The polar average reference effect: a bias in estimating the head surface integral in EEG recording. Clin. Neurophysiol. 110 (6), 1149–1155. https://doi.org/10.1016/S1388-2457(99)00044-9.

King, A., Hopkins, K., Plack, C.J., 2016. Differential group delay of the frequency following response measured vertically and horizontally. J. Assoc. Res. Otolaryngol. 17 (2), 133–143. https://doi.org/10.1007/s10162-016-0556-x.

Krishnan, A., Xu, Y., Gandour, J.T., Cariani, P.A., 2004. Human frequency-following response: representation of pitch contours in Chinese tones. Hear. Res. 189 (1–2), 1–12. https://doi.org/10.1016/S0378-5955(03)00402-7.

Krizman, J., Kraus, N., 2019. Analyzing the FFR: a tutorial for decoding the richness of auditory function. Hear. Res. 382, 107779 https://doi.org/10.1016/j.heares.2019.107779.

Kulasingham, J.P., Brodbeck, C., Presacco, A., Kuchinsky, S.E., Anderson, S., Simon, J.Z., 2020. High gamma cortical processing of continuous speech in younger and older listeners. Neuroimage 222, 117291. https://doi.org/10.1016/j.neuroimage.2020.117291.

Lau, B.K., Ruggles, D.R., Katyal, S., Engel, S.A., Oxenham, A.J., 2017. Sustained cortical and subcortical measures of auditory and visual plasticity following short-term perceptual learning. PLoS One 12 (1), e0168858. https://doi.org/10.1371/journal.pone.0168858.

Lenth, R.v., 2016. Least-squares means: the R package lsmeans. J. Stat. Software 69 (1), 1–33. https://doi.org/10.18637/jss.v069.i01.

Lu, H., Mehta, A.H., Bharadwaj, H.M., Shinn-Cunningham, B.G., Oxenham, A.J., 2020. Comment on 'Rapid acquisition of auditory subcortical steady state responses using multichannel recordings. Clin. Neurophysiol. 131 (8), 1833–1834. https://doi.org/10.1016/j.clinph.2020.05.018.

Luke, R., Wouters, J., 2017. Kalman filter based estimation of auditory steady state response parameters. IEEE Trans. Neural Syst. Rehabil. Eng. 25 (3), 196–204. https://doi.org/10.1109/TNSRE.2016.2551302.

Mao, D., Innes-Brown, H., Petoe, M.A., Wong, Y.T., McKay, C.M., 2018. Cortical auditory evoked potential time-frequency growth functions for fully objective hearing threshold estimation. Hear. Res. 370, 74–83. https://doi.org/10.1016/j.heares.2018.09.006.

Mehta, A.H., Feng, L., Oxenham, A.J., 2021. Neural auditory contrast enhancement in humans. Proc. Natl. Acad. Sci. USA 118 (29), e2024794118. https://doi.org/10.1073/pnas.2024794118.

Mijares, E., Pérez Abalo, M.C., Herrera, D., Lage, A., Vega, M., 2013. Comparing statistics for objective detection of transient and steady-state evoked responses in newborns. Int. J. Audiol. 52 (1), 44–49. https://doi.org/10.3109/14992027.2012.736030.

Parthasarathy, A., Bartlett, E., 2012. Two-channel recording of auditory-evoked potentials to detect age-related deficits in temporal processing. Hear. Res. 289 (1–2), 52–62. https://doi.org/10.1016/j.heares.2012.04.014.

Paul, B.T., Bruce, I.C., Roberts, L.E., 2017. Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus. Hear. Res. 344, 170–182. https://doi.org/10.1016/j.heares.2016.11.010.

Picton, T.W., John, M.S., Dimitrijevic, A., Purcell, D., 2003. Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. Int. J. Audiol. 42 (4), 177–219. https://doi.org/10.3109/14992020309101316.

Picton, T.W., Vajsar, J., Rodriguez, R., Campbell, K.B., 1987. Reliability estimates for steady-state evoked potentials. Electroencephalogr. Clin. Neurophysiol. Evoked Potentials 68 (2), 119–131. https://doi.org/10.1016/0168-5597(87)90039-6.

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., Authors, E., R-core, 2019. Linear and non-linear mixed effects models. Package "nlme", version 3, 1–141. Comprehensive R Archive Network (CRAN). https://cran.r-project.org/web/packages/nlme/index.html.

Polonenko, M.J., Maddox, R.K., 2019. The parallel auditory brainstem response. Trends in Hearing 23, 233121651987139. https://doi.org/10.1177/2331216519871395.

Poulsen, C., Picton, T.W., Paus, T., 2009. Age-related changes in transient and oscillatory brain responses to auditory stimulation during early adolescence. Dev. Sci. 12 (2), 220–235. https://doi.org/10.1111/j.1467-7687.2008.00760.x.

Rance, G., 2008. The Auditory Steady-State Response: Generation, Recording, and Clinical Application. Plural Publishing.

Ross, B., Herdman, A.T., Pantev, C., 2005. Right hemispheric laterality of human 40 Hz auditory steady-state responses. Cerebr. Cortex 15 (12), 2029–2039. https://doi.org/10.1093/cercor/bhi078.

Samuelsson, J.G., Khan, S., Sundaram, P., Peled, N., Hämäläinen, M.S., 2019. Cortical signal suppression (CSS) for detection of subcortical activity using MEG and EEG. Brain Topogr. 32 (2), 215–228. https://doi.org/10.1007/s10548-018-00694-5.

Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a tutorial. Ear Hear. 31 (3), 302–324. https://doi.org/10.1097/AUD.0b013e3181cdb272.

Skosnik, P.D., Krishnan, G.P., O'Donnell, B.F., O'Donnell, B.F., 2007. The effect of selective attention on the gamma-band auditory steady-state response. Neurosci. Lett. 420 (3), 223–228. https://doi.org/10.1016/J.NEULET.2007.04.072.

Sohmer, H., Pratt, H., Kinarti, R., 1977. Sources of frequency following responses (FFR) in man. Electroencephalogr. Clin. Neurophysiol. 42 (5), 656–664. https://doi.org/10.1016/0013-4694(77)90282-6.

Spencer, K.M., 2012. Baseline gamma power during auditory steady-state stimulation in schizophrenia. JANUARY 2012 Front. Hum. Neurosci. 5, 1–7. https://doi.org/10.3389/fnhum.2011.00190.

Spencer, K.M., Salisbury, D.F., Shenton, M.E., McCarley, R.W., 2008. Gamma-band auditory steady-state responses are impaired in first episode psychosis. Biol. Psychiatr. 64 (5), 369–375. https://doi.org/10.1016/j.biopsych.2008.02.021.

Stapells, D.R., Makeig, S., Galambos, R., 1987. Auditory steady-state responses: threshold prediction using phase coherence. Electroencephalogr. Clin. Neurophysiol. 67 (3), 260–270. https://doi.org/10.1016/0013-4694(87)90024-1.

Swaminathan, J., Krishnan, A., Gandour, J.T., 2008. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. Neuroreport 19 (11), 1163–1167. https://doi.org/10.1097/WNR.0b013e3283088d31.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J. Neurosci. 16 (13), 4240–4249. https://doi.org/10.1523/jneurosci.16-13-04240.1996.

Tan, H.-R.M., Gross, J., Uhlhaas, P.J., 2015. MEG—measured auditory steady-state oscillations show high test–retest reliability: a sensor and source-space analysis. Neuroimage 122, 417–426. https://doi.org/10.1016/j.neuroimage.2015.07.055.

Team, R Core, 2018. R: A Language and Environment for Statistical Computing. https://www.r-project.org/.

Thatcher, R.W., 2012. Coherence, phase differences, phase shift, and phase lock in EEG/ERP analyses. Dev. Neuropsychol. 37 (6), 476–496. https://doi.org/10.1080/87565641.2011.619241.

Tichko, P., Skoe, E., 2017. Frequency-dependent fine structure in the frequency-following response: the byproduct of multiple generators. Hear. Res. 348, 1–15. https://doi.org/10.1016/j.heares.2017.01.014.

van Dun, B., Wouters, J., Moonen, M., 2007. Improving auditory steady-state response detection using independent component analysis on multichannel EEG data. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 54 (7), 1220–1230. https://doi.org/10.1109/TBME.2007.897327.

van Dun, B., Wouters, J., Moonen, M., 2009. Optimal electrode selection for multi-channel electroencephalogram based detection of auditory steady-state responses. J. Acoust. Soc. Am. 126 (1), 254–268. https://doi.org/10.1121/1.3133872.

Vander Werff, K.R., 2009. Accuracy and time efficiency of two ASSR analysis methods using clinical test protocols. J. Am. Acad. Audiol. 20 (7), 433–452. https://doi.org/10.3766/jaaa.20.7.5.

Vanheusden, F.J., Bell, S.L., Chesnaye, M.A., Simpson, D.M., 2019. Improved detection of vowel envelope frequency following responses using Hotelling's T2 analysis. Ear Hear. 40 (1), 116–127. https://doi.org/10.1097/AUD.0000000000000598.

Varghese, L., Bharadwaj, H.M., Shinn-Cunningham, B.G., 2015. Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. Brain Res. 1626, 146–164. https://doi.org/10.1016/j.brainres.2015.06.038. Elsevier.

White-Schwoch, T., Anderson, S., Krizman, J., Nicol, T., Kraus, N., 2019. Case studies in neuroscience: subcortical origins of the frequency-following response. J. Neurophysiol. 122 (2), 844–848. https://doi.org/10.1152/JN.00112.2019.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Syst. 2 (1–3), 37–52. https://doi.org/10.1016/0169-7439(87)80084-9.

Zhang, X., Gong, Q., 2017. Correlation between the frequency difference limen and an index based on principal component analysis of the frequency-following response of normal hearing listeners. Hear. Res. 344, 255–264. https://doi.org/10.1016/j.heares.2016.12.004.

Zhu, L., Bharadwaj, H., Xia, J., Shinn-Cunningham, B., 2013. A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. J. Acoust. Soc. Am. 134 (1), 384–395. https://doi.org/10.1121/1.4807498.