# A Novel EEG Paradigm to Simultaneously and Rapidly Assess the Functioning of Auditory and Visual Pathways

Kristina C. Backer[1,2]*, Andrew S. Kessler[1]*, Laurel A. Lawyer[1], David P. Corina[1,3], and Lee M. Miller[1,4]

[1] Center for Mind and Brain, University of California, Davis, USA

[2] Dept. of Cognitive and Information Sciences, University of California, Merced, USA

[3] Dept. of Linguistics, University of California, Davis, USA

[4] Dept. of Neurobiology, Physiology, and Behavior, University of California, Davis, USA

* These authors contributed equally to this work.

**Author Contributions**

ASK, LAL, DPC, and LMM conceived and designed the research. ASK and LAL performed experiments. KCB and ASK analyzed data. All authors interpreted results. KCB prepared the figures and drafted the manuscript. All authors edited and revised the manuscript and approved the final version of the manuscript.

**Running Head**

A Novel Auditory-Visual EEG Paradigm

**Address for Correspondence**

Lee M. Miller

UC Davis Center for Mind and Brain

267 Cousteau Place

Davis, CA 95618

Email: leemiller@ucdavis.edu

Phone: (530) 754-4574

## Abstract

Objective assessment of the sensory pathways is crucial for understanding their development across the lifespan and how they may be affected by neurodevelopmental disorders (e.g., autism) and neurological pathologies (e.g., stroke, multiple sclerosis, etc.). Quick and passive measurements, for example using electroencephalography (EEG), are especially important when working with infants and young children, and with patient populations having communication deficits (e.g., aphasia). However, many EEG paradigms are limited to measuring activity from one sensory domain at a time, may be time consuming, and target only a subset of possible responses from that particular sensory domain (e.g., only auditory brainstem responses or only auditory P1-N1-P2 evoked potentials). Thus, we developed a new multisensory paradigm that enables simultaneous, robust, and rapid (6-12 minute) measurements of both auditory and visual EEG activity, including auditory brainstem responses (ABRs), auditory and visual evoked potentials, as well as auditory and visual steady-state responses. This novel method allows us to examine neural activity at various stations along the auditory and visual hierarchies with an ecologically valid continuous speech stimulus, while an unrelated video is playing. Both the speech stimulus and the video can be customized for any population of interest. Furthermore, by using two simultaneous visual steady-state stimulation rates, we demonstrate the ability of this paradigm to track both parafoveal and peripheral visual processing concurrently. We report results from twenty-five healthy young adults, which validate this new paradigm.



**Keywords**: auditory, visual, evoked potentials, steady-state responses, Cheech

**New and Noteworthy**

A novel electroencephalography (EEG) paradigm enables the rapid, reliable, and non-invasive assessment of neural activity along both auditory and visual pathways concurrently. The paradigm uses an ecologically valid continuous speech stimulus for auditory evaluation and can simultaneously track visual activity to both parafoveal and peripheral visual space. This new methodology may be particularly appealing to researchers and clinicians working with infants and young children, and with patient populations with limited communication abilities.

## Introduction

To understand sensory development across the lifespan and the impact of neurodevelopmental disorders (e.g., autism spectrum) or neurological pathologies and insults (e.g., multiple sclerosis, stroke, etc.) on sensory systems, the ability to objectively measure the functioning of sensory pathways is critical. Reliable objective and passive measures are especially important when working with individuals with limited communication abilities (e.g., infants, individuals with aphasia, etc.). Furthermore, from a research and potentially a clinical standpoint, the ability to objectively, non-invasively, and quickly assess the functioning of visual and auditory pathways can provide important information about an individual that is not readily available through behavioral testing. For instance, this information may be used to link individual differences in a child's sensory development with his or her cognitive development, or to guide research and development of individualized clinical interventions. In our case, we developed the paradigm described herein to examine sensory development in normal-hearing children and children with cochlear implants. In the current manuscript, we present data from healthy young adults as validation of the methodology.

The objective of the current paradigm was to record numerous clinically important auditory and visual neural responses simultaneously and quickly, while the participant watched an unrelated video. While several non-invasive neuroimaging techniques could be used to achieve this goal, we chose electroencephalography (EEG) for its many practical advantages. EEG has excellent (sub-millisecond) temporal resolution, which is essential to examine neural activity along the auditory hierarchy from the brainstem to the cortex and to track auditory and visual steady-state responses. Furthermore, EEG is safe and has been used for decades in clinical settings. It is portable and relatively inexpensive to use, unlike magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI). In addition, EEG poses no contra-indications unlike fMRI, with which many metallic medical

devices or implants may be incompatible due to safety risks. Finally, EEG can measure activity from deep brain structures, particularly the auditory brainstem, unlike functional near-infrared spectroscopy (fNIRS) which is limited to superficial cortical regions. Thus, for the purposes of our objective, the above-mentioned strengths of EEG outweighed its primary weakness, that is, poor spatial resolution.

Many EEG protocols are limited to collecting a subset of neural responses in one sensory modality at a time. Furthermore, especially in the auditory modality, there has been increased interest in assessing relationships between the early brainstem EEG activity to the later cortical responses, and eventually to speech perception, within individuals. Paradigms that have been designed for this purpose are either time consuming (due to the longer inter-stimulus interval needed for the later responses juxtaposed with the large number of sweeps required for reliable early brainstem responses, if recording both types of responses simultaneously) or are unable to record the early and later responses simultaneously (e.g., Bidelman 2015; Bidelman et al. 2013; Krishnan et al. 2012; Musacchia et al. 2008; Woods et al. 1993). A paradigm recently developed by Slugocki et al. (2017) simultaneously measured both subcortical and cortical responses (including P3a and mismatch negativity (MMN)) to auditory stimuli; however, the recording time was relatively long (approx. 40 minutes) and the stimuli were created using amplitude-modulated tones. Another group developed an EEG paradigm to simultaneously record potentials including the auditory N1, MMN, P300, and N400 in about 5 minutes, using both tone and speech stimuli (Sculthorpe-Petley et al. 2015); however, this paradigm was limited to cortical potentials in response to auditory stimuli only. In contrast, the EEG paradigm described herein, allows for the rapid recording of both auditory and visual responses simultaneously, using a specially engineered continuous speech stimulus and an interspersed visual stimulus. The continuous speech stimulus allows for the examination of auditory EEG activity from brainstem to cortex and under more naturalistic conditions, compared to other

frequently used stimuli like clicks, tones, and consonant-vowel syllables. The visual stimulus permits assessment of transient evoked and steady-state visual responses.

It is important to mention that this paradigm significantly builds upon previous work conducted by our group (Miller et al. 2017). Previously, a continuous speech stimulus was used and demonstrated the feasibility of simultaneously obtaining auditory evoked responses along the auditory pathway, from the brainstem to the cortex. This set the foundation for the current methodology. Novel aspects of the EEG paradigm presented in this report include the following: First, in conjunction with the continuous speech stimulus, unrelated visual flicker stimuli were used to obtain both auditory and visual EEG responses simultaneously. Second, in the current implementation, the use of a silent video, which engages attention and is unrelated to the auditory and visual stimuli, makes the paradigm suitable to different populations (e.g., young children). This is a critical validation step, to determine which responses can be reliably observed – even when the auditory and visual stimuli are not necessarily attended.

To introduce the current methodology, we provide an overview of the auditory and visual responses that the EEG paradigm was designed to measure, including the time course for each response, how the response can be elicited, and the putative neural generators for each response. The following sections therefore illustrate the scope of our approach and motivate many technical details of the design and analysis, described next. It also serves as a brief tutorial for readers unfamiliar with auditory and visual EEG. (For a thorough discussion of these topics, see Halgren 1990; Hall 2007; Luck 2014.)

**Auditory Responses**

The stimulus used in the current EEG paradigm was designed to enable the simultaneous recording of the auditory brainstem response (ABR), the middle latency response

(MLR), the long latency response (LLR), as well as the auditory steady-state response (ASSR), all in the context of naturalistic, intelligible, and continuous spoken language.

The stereotyped ABR consists of seven positive peaks (Waves I to VII) that occur within 10 ms following the onset of a brief sound; Wave V generally has the largest amplitude of these peaks (Jewett and Williston 1971). Often, the ABR is elicited with a click stimulus (e.g., Jewett and Williston 1971; Pratt and Sohmer 1976), but tone pips (e.g., Suzuki et al. 1977; Weber and Folsom 1977; Woldorff and Hillyard 1991), chirps (e.g., Bell et al. 2002a; Dau et al. 2000; Elberling and Don 2008), and brief speech sounds (e.g., consonant-vowel syllables, Krizman et al. 2010) can elicit ABRs as well. The ABR reflects the neural response to sound ascending the auditory pathway (for a review, see Moore 1987), from the eighth cranial nerve (Wave I) (Hashimoto et al. 1979; Moller et al. 1982; Moller et al. 1981; Starr and Hamilton 1976) to the lateral lemniscus and inferior colliculus (Waves IV and V) (Moore 1987; Starr and Hamilton 1976).

The MLR is the next set of waveforms as the acoustic representation ascends along the auditory pathway. The MLR comprises two negative peaks interleaved with two positive peaks (Na, Pa, Nb, and Pb), which occur from approximately 15 to 60 ms after sound onset (Geisler et al. 1958; Goldstein and Rodman 1967). Some studies have also reported waves $N_0$ and $P_0$, which occur earlier, around 8-9 and 12-14 ms, respectively (Mendel and Goldstein 1969; Picton et al. 1974; Yoshiura et al. 1996). Clicks, tones, and chirps can elicit the MLR (e.g., Bell et al. 2002b; Mendel and Goldstein 1969; Picton et al. 1974). Taken together, source modeling (Pelizzone et al. 1987; Rupp et al. 2002; Scherg and Von Cramon 1986; Yoshiura et al. 1996; 1995), intracranial (Celesia 1976; Lee et al. 1984; Liegeois-Chauvel et al. 1994), and lesion (Kileny et al. 1987; Kraus et al. 1982) studies have shown that the MLR is primarily generated in supratemporal cortex. Additionally, sub-cortical activity likely contributes to at least the earlier MLR components, especially the Na (Hashimoto 1982; Kileny et al. 1987).

The LLR is the final set of auditory evoked potentials observed in the cascade. A stereotyped LLR includes the P1[1], N1, P2, and N2 components, which typically span approximately 50 to 300 ms, following sound onset (Davis and Zerlin 1966; Davis 1939; Vaughan and Ritter 1970). A variety of sounds, including clicks (e.g., Arslan et al. 1984), tones (e.g., Davis and Zerlin 1966), and speech sounds (e.g., Kraus et al. 1993), can be used to elicit the LLR. The primary and non-primary auditory cortices, with contributions from the association and possibly frontal cortices, are the putative generators of the LLR (Hari et al. 1980; Kanno et al. 2000; Naatanen and Picton 1987; Picton et al. 1999; Scherg and Von Cramon 1985; Shahin et al. 2007; Vaughan and Ritter 1970).

A fourth measure of auditory function that our paradigm was designed to elicit is the ASSR (also known as the 40 Hz Response) (Galambos et al. 1981). Transient stimuli (e.g., clicks or tones) that repeat at a constant rate (often ~40 Hz, or every ~25 ms, in the auditory domain), or amplitude-modulated tones or noise can lead to a sinusoidal-shaped event-related potential (ERP), also known as a steady-state response (for a review, see Korczak et al. 2012; Picton et al. 2003)[2]. Generally, the ASSR is analyzed in the frequency domain (via a Fourier transform) or time-frequency domain (e.g., via wavelet decomposition), so that amplitude peaks at the stimulation frequency and its harmonics are clearly visible (e.g., Artieda et al. 2004; Stapells et al. 1984). Regarding the neural generators of the ASSR, taken together, MEG and EEG studies have demonstrated that both the auditory brainstem and auditory cortex contribute to the 40-Hz ASSR (e.g., Makela and Hari 1987; Ross et al. 2002; Schoonhoven et al. 2003; Coffey et al. 2016; Herdman et al. 2002).

**Visual Responses**

The current paradigm was also designed to elicit onset visual evoked potentials (VEPs) and the steady-state visual evoked potential (SSVEP). The onset VEP generally consists of the

P1, N1, and P2 components, which can be evoked by a variety of visual stimuli, including flashes, checkerboard, or grating stimuli. The VEP complex is evident from approximately 50 to 250 ms after visual stimulus onset (e.g., Clark et al. 1994; Jeffreys and Axford 1972). Many studies have localized the P1 and N1 to extrastriate regions (e.g., Clark et al. 1994; Di Russo et al. 2002; Gomez Gonzalez et al. 1994), but both striate and extrastriate regions may contribute, at least to the P1 (Aine et al. 1995; Di Russo et al. 2005; Vanni et al. 2004). The generators of the visual P2 are not well understood and likely involve multiple cortical sources (Clark et al. 1994). However, the P2 has been shown to peak over the vertex following both auditory and visual stimuli, suggesting that neurons in amodal cortical regions may contribute to both auditory and visual P2 responses (Perrault and Picton 1984).

Unlike onset VEPs, which are transient onset responses, SSVEPs are brain responses to a flickering visual stimulus that has a constant flicker rate, such as sinusoidally-modulated flashes of light (Regan 1966; Van Der Tweel and Lunel 1965) or checkerboards in which the black checks change to white and vice versa at a constant rate (e.g., Burkitt et al. 2000; Thorpe et al. 2007). SSVEPs have a spectral amplitude distribution that remains stable over time and reflects the visual stimulus' flicker rate (and its harmonics); thus, instead of analyzing SSVEPs in the time domain, they are generally analyzed in the frequency domain, using a Fourier transform, or in the time-frequency domain (e.g., using a wavelet approach) (for reviews, see Norcia et al. 2015; Regan 1977; Vialatte et al. 2010). One observation critical to the design of the current paradigm is that multiple visual stimuli with different flicker rates can be presented concurrently to "frequency-tag" neural activity, in other words, to isolate SSVEPs for each separate stimulus/rate presented (e.g., Andersen et al. 2008; Ding et al. 2006; Itthipuripat et al. 2013; Keitel et al. 2010; Muller et al. 2003; Regan and Heron 1969).

**Expected Findings.**

We anticipated that the ABR, MLR, LLR, and ASSR, as well as the VEP and SSVEP, would be reliably detected in all individuals. Previous studies which used long-duration auditory stimuli (Krishnan et al. 2012; Picton et al. 1978a; b) showed that the resulting LLR does not have canonical P1-N1-P2 morphology, but rather it has a broad P1, followed by an N1 and a sustained negativity. Thus, since the present paradigm employs continuous speech stimuli, we expected similar non-canonical morphology as shown in the aforementioned studies.

## Materials and Methods

### Participants

Twenty-six healthy young adult volunteers participated in this study; however, due to technical problems during one participant's session, data from 25 participants were analyzed (14 females, Age Range: 18 to 28 years, Mean Age: 21 years; one participant did not provide his age). Participants were right-handed, fluent English speakers, who self-reported normal hearing, normal or corrected-to-normal visual acuity, normal color vision, no history of any neurological illnesses, and no known problems with speech or reading. All participants gave informed written consent before commencing the study, and the UC Davis Institutional Review Board approved all procedures described herein.

### Visual Stimuli

As illustrated in **Figure 1A**, the visual stimuli included a cartoon played in the middle of the screen, surrounded by two concentric checkered rings. The cartoon serves as an engaging event and is designed to help control a participant's fixation. For this study, we used a variety of cartoon clips and videos of animals that were gleaned from the internet and compiled and edited using Final Cut Pro software (Apple, Inc.). These video clips were selected for subsequent

testing of elementary school-aged children. We took care to avoid or edit out portions of cartoons which involved characters talking, to prevent any attempts to audio-visually integrate the cartoon and the Cheech. The inner ring comprised eight equally-spaced checks, which flickered sinusoidally at a rate of 7.5 Hz. The outer ring comprised 16 equally- spaced checks, which flickered sinusoidally at 12 Hz. To prevent the flickering from being overly bothersome to participants, the rings flickered for 2.5 seconds, and then stopped for 1-3 seconds (randomly jittered, rectangular distribution), before the flickering began again. In each ring, alternate checks flickered in counter-phase, as depicted in **Figure 1A**. Additionally, within each ring and between every flickering check was a "blank" check or gap, the same color as the background, which did not flicker; this was done to prevent multiple checks stimulating a given location on the retina, as eye movements naturally occur during cartoon viewing. In this way, since no adjacent checks flicker out of phase, no retinal location receives phase-opposed stimulation over time during small eye movements.

The visual stimuli were created using MATLAB (The MathWorks, Inc., https://www.mathworks.com/), and the cartoon and flickering rings were embedded into six 2-minute videos (AVI files). Participants sat 32 inches (~81.28 cm) from the monitor. The spacing between checks (angular distance), as well as the radial spacing between rings was chosen such that it was approximately half the cartoon width – again to avoid multiple checks stimulating the same retinal location. The cartoon extended ~2.7° (visual angle) to the left and right of screen center (~5.5° total cartoon width), and ~1.8° (visual angle) above and below screen center (~3.66° total cartoon height). The inner edge of the inner ring was adjacent to the border of the cartoon (see **Figure 1**), and the inner ring's outer edge subtended a visual angle of ~5.8° from screen center, while the inner edge of the outer ring was ~10.1° from screen center. The outer ring was made larger than the inner ring to approximately account for cortical magnification (Cowey and Rolls 1974; Daniel and Whitteridge 1961). The individual checks of

the outer ring extended to varying visual angles from screen center, according to the boundaries of the screen edge. For example, at the corners of the screen (longest distance from screen center), the outer ring extended to a visual angle of ~20.5°, whereas at the shortest distance from screen center (top or bottom of the screen, directly above or below screen center, respectively), the outer ring extended to a visual angle of ~11.3°. The cartoon was rendered at 147 x 98 pixels and combined in MATLAB with the flickering ring stimuli which extended 960 x 600 pixels. Upon stimulus delivery via Presentation software (Neurobehavioral Systems, Inc., https://neurobs.com), which doubled the video size, the cartoon resolution was 294 x 196 pixels and the full video resolution was 1920 x 1200 pixels.

We originally manipulated the color/luminance of the checks to determine the paradigm's sensitivity to changes in color/luminance. In one block (three 2-minute videos per block), the checks were black and white against a gray background (RGB color components: [0 0 0]; [255 255 255]; [128 128 128], respectively); in the other block, the checks were red and green, against a mustard yellow background (RGB color components: [255 0 0]; [0 255 0]; [128 128 0], respectively). Block order was counterbalanced across participants. The flickering involved luminance changes between black and white or color changes between red and green (the gray and yellow intervening gaps did not change). Rather than using abrupt luminance or color transitions, a sinusoidal function was applied, such that the speed of the transition depended on the screen's refresh rate (60 Hz) and the flicker rate. Thus, for the inner ring (7.5 Hz flicker), the color/luminance transition occurred gradually across eight frames (i.e., screen refreshes), while for the outer ring (12 Hz flicker), the transition occurred across five frames. During each block, there were 90 2.5-second intervals, in which the checkered rings were flickering. The EEG data were time-locked to these flicker onsets to compute the visual onset responses and SSVEPs. The two visual color conditions (black-white, red-green) were not isoluminant, thereby

confounding interpretations about any differences observed between color conditions. Thus, we collapsed across black-white and red-green trials for EEG data analysis.

**Auditory Stimuli**

**Figure 1B** shows an example of the auditory stimuli. A detailed characterization of these auditory stimuli (termed CHirp-spEECH, “Cheech”) can be found in the patent listing (Miller et al. 2017). Unlike fully natural speech, Cheech possesses acoustic properties that robustly drive early (ABR) as well as middle and late auditory EEG responses. As implied in its name, Cheech incorporates auditory chirp stimuli; chirps are transient sounds that increase rapidly in frequency over time. Furthermore, Cheech takes advantage of the observation that upward frequency-modulated chirps yield more synchronized brainstem responses than traditional stimuli such as clicks, by compensating for the traveling wave delay (across frequencies) along the basilar membrane (Dau et al. 2000; Elberling et al. 2007; Shore and Nuttall 1985). In Cheech, we replace some of the glottal pulse energy with chirp energy, thereby yielding stronger speech EEG responses. Briefly, 49 unique sentences (sampled at 22,050 Hz) from the Harvard/IEEE Corpus (1969) were selected and concatenated into a two-minute WAV file of continuous speech. Next, the pitch of the voicing was flattened to 82 Hz using Praat (Boersma and Weenink 2001, http://www.praat.org/). A second sound was then created with trains of chirps temporally coinciding with each voiced period, such that the individual chirps were aligned in time with individual glottal pulses in the speech (i.e., every other glottal pulse). Voiced periods with coinciding chirps occurred whenever the speech envelope <40Hz for energy between 20-1000Hz (containing the highest voiced power) surpassed a threshold (appx. 28% of overall speech RMS) long enough to contain four chirps total. Finally, the speech and chirps were frequency multiplexed in alternating, interleaved bands one octave wide and added together (speech energy occupied [0 250], [500-1000], [2000-4000] Hz, and chirp energy occupied [250-500], [1000-2000], [4000 10,000] Hz). In this

way, chirps align acoustically and perceptually with the natural voicing, creating a single perceptual speech object. Chirps occurred at a rate of 41 Hz within each voiced period, to elicit the ASSR at 41 Hz. Furthermore, the chirps were isochronous throughout the WAV file (due to the flattened pitch), so that each chirp occurred at multiples of ~24 ms, relative to the first chirp in each experimental block. Within each voiced period, the second chirp was omitted, in order to measure an MLR that occurred in response to the onset of the voiced period (which coincides with the first chirp). The voicing periods occurred on average 501 ms apart (range: 146 to 1195 ms apart) in the present study. The resulting stimulus is highly intelligible speech, albeit with a robotic monotone quality; the rapid interspersed chirps are audible as a rattling character in the voicing, but they blend perceptually with the speech and do not distract from its linguistic content.

Across both 6-minute blocks, there were 1,422 voicing onsets (711 per block) and 11,694 chirps (5,847 per block). As we demonstrate here, using Cheech in conjunction with EEG, robust measurements of auditory responses along the entire auditory pathway, from the brainstem to auditory cortex, can be obtained (Miller et al. 2017). Specifically, ABRs can be created by time-locking and signal averaging the continuous EEG data, relative to the onset of each chirp. By time-locking the EEG data to the onsets of the voicing periods, MLRs, LLRs, and ASSRs can be measured.

In this study, the two-minute WAV file of Cheech was repeated three times within each of the two blocks. The Cheech was presented in the free-field at a level of 65 dB(C) SPL, using an internal Realtek HD sound card, a NuForce Icon stereo amplifier, and finally played through an Auvio 400016 speaker (passive speaker, three drivers in speaker chassis) that was positioned 1.27 m directly in front of the participant and above the computer screen. Previously, Cheech has been presented monaurally (Miller et al. 2017), but we chose to use the free-field approach because of our intention to use this paradigm in children with cochlear implants. A

sample stimulus video, which includes the Cheech audio, can be accessed online here: https://figshare.com/articles/Novel_EEG_Paradigm_J_Neurophysiology_Methods_Backer_et_al_2019pptx_pptx/8214449.

**Electroencephalography (EEG) Recording**

EEG data were recorded using a BioSemi ActiveTwo system (BioSemi B.V., Netherlands, biosemi.com), a 32-channel cap, and ActiView2 software installed on a Dell laptop. The scalp electrode montage (based on the International 10/20 System) included: FP1/2, AF3/4, Fz/3/4/7/8, FC1/2/5/6, Cz/3/4, T7/8, CP1/2/5/6, Pz/3/4/7/8, PO3/4, Oz/1/2. Additional electrodes were taped to each earlobe and each mastoid. EEG data were sampled at a rate of 16,384 Hz in order to obtain ABRs, which require sub-millisecond resolution; an anti-aliasing low-pass filter at 3,334 Hz ($5^{th}$ order sinc) was applied before A-to-D conversion. Before beginning the recording, electrode offsets (relative to the Common Mode Sense (CMS) electrode) for all channels were set to < 20 μV.

During EEG recording, the visual and auditory (Cheech) stimuli were simultaneously presented using Presentation software, from a Dell laptop to a 24-inch HP Z Display monitor and to the speaker inside the testing room. The EEG recording lasted for a total time of 12 minutes (6 minutes per block). Participants were instructed to focus on the cartoon, with no explicit instruction to ignore the flickering visual or the Cheech stimuli. As mentioned previously, the inter-stimulus interval for the visual flicker stimuli was jittered from one to three seconds, while the cartoon was played continuously; however, the Cheech was looped without any silent periods. Thus, there were periods within each block, in which the participants experienced only Cheech and no visual flickers (auditory-only); at other times, both the flickers and the Cheech were concurrent (audio-visual) (**Figure 1C**). All auditory events (i.e., during both the auditory-

only and audio-visual periods) were used in computing the auditory EEG responses reported here.

**EEG Data Analysis**

***Preprocessing.***

EEG data were preprocessed in MATLAB, using EEGLAB (Delorme and Makeig 2004), ERPLAB Toolbox (Lopez-Calderon and Luck 2014), and custom MATLAB code. First, raw data (BDF files) were imported into EEGLAB using the BioSig plugin (version 2.88, https://sourceforge.net/projects/biosig/).

During EEG acquisition, experimental events were synchronized with the EEG data using parallel port codes, sent from the presentation computer, using a StarTech IEEE 1284 parallel port card, to the Biosemi acquisition box. For the visual stimuli, a single port code was sent via Presentation software at the onset of each 2-minute video; but the time stamp corresponding to the onsets of the flicker interval within each video were added post-hoc, relative to each video onset, using custom MATLAB code. Since the videos were created in MATLAB, the exact frames corresponding to the start of the flicker interval were known in advance. Using Presentation's detailed logging feature for videos, we obtained information about the onset time for each frame, the uncertainty about each frame's onset time (usually 1-2 ms), and the number of frames dropped for each participant (usually no frames dropped). Furthermore, using a photodiode and oscilloscope, we checked the timing between the delivery of the EEG port code at the start of each video and the actual start of the video to ensure reliable timing.

For the auditory stimuli, Presentation sent one port code at the beginning of each WAV file, as well as port codes for all of the Cheech events of interest (i.e., voicing onsets and chirps). These were all embedded as metadata within the WAV files and were sent by

Presentation at appropriate latencies as the WAV files played. It is important, particularly for ABR analyses, to ensure accurate sub-millisecond precision between the port code times and when the Cheech events of interest actually played; however, in our system, comparing the sound output and port output using an oscilloscope, the port code timing variability was too great for ABR analyses (~1 ms jitter). In addition, port codes were frequently missed due to the unusually high load for the parallel port in sending codes every ~24 ms. Many acquisition systems would not suffer these limitations, as they provide an additional data channel dedicated to timing pulses aligned precisely with the stimuli (further described in the Discussion). However, in our own system, a crucial step involved developing MATLAB code to correct these timing issues, particularly for the ABRs. Thus, we conducted a cross-covariance analysis between the recorded port codes in the EEG data and the intended port codes embedded in the WAV file metadata (taking into account the difference in sampling rates between the WAV files and EEG recordings). This yields large covariance peaks at the onset of each WAV file, which we then used as a temporal reference to replace all recorded port codes with reconstructed ones from the WAV files themselves. Thus, the event timing used in the analyses has essentially no variability due to port code timing errors. An analysis of the timing difference between the recorded port codes and the reconstructed port codes revealed that 8.95% of the original, recorded auditory port codes were jittered by more than 0.2 ms.

One further temporal correction was necessary, due to the fact that experimental devices may differ slightly in how they measure time. Thus, over long recording periods, the stimulus presentation computer or sound card can nominally drift out of sync from the EEG acquisition device. Put in another way, reported time runs slightly slower or faster for different devices. This required us to compress time for the target port codes by a very small percentage (0.003%) to accommodate the difference.

Due to divergent analyses after the addition/correction of port codes, the various preprocessing pipelines are described separately below.

Auditory LLRs and Visual Responses.

Following port code addition/correction, the data were resampled to 512 Hz, and each block's data was concatenated. Next, the data were inspected visually across the entire recording, noisy data segments were removed, and bad channels were noted. Each participant's data were referenced to the average earlobes, and band-pass filtered from 0.5 to 100 Hz, using an $8^{th}$ order, zero-phase Butterworth filter. The DC offset of contiguous segments of data was removed prior to filtering, in order to minimize edge effects at boundaries. One participant's data was also filtered using a Park's-McClellan notch filter (order of 180) to remove 60 Hz noise. The filtered data were then processed with EEGLAB's Independent Component Analysis (ICA) function, which used the Infomax algorithm (Bell and Sejnowski 1995); the reference (earlobe) channels, as well as any bad channels, were excluded from ICA. The components were visually inspected, and only eye blink components were removed from the data; twenty-three participants had one component removed, and two participants had two components removed. For the data used to create visual onset ERPs, as well as the auditory long latency responses, the data were down-sampled further to 256 Hz. Next, any bad channels identified previously were spatially interpolated using a spline function; ten participants had no bad channels, six had one bad channel, five had two bad channels, two participants had three bad channels, and two participants had five bad channels.

*Filtering and Epoching.* The next steps occurred in the same order for each participant's data, however, the filter and epoch settings differed depending on the type of response that was being extracted, as detailed below. Following interpolation of bad channels, the data were filtered with an appropriate zero-phase, $8^{th}$ order Butterworth filter to obtain the desired

passband for each response type. For the VEP onset response and the auditory LLR, the data were low-pass filtered with a 30 Hz cutoff frequency, resulting in a passband of 0.5 to 30 Hz. The SSVEP was analyzed in the frequency domain (using a Fourier transform). For the SSVEP response, a band-pass filter was applied with cutoffs at 1 and 40 Hz. Next, using ERPLAB Toolbox, information about the time-locking events of interest (e.g., flickering onsets) was obtained, and the data were epoched and baselined to the pre-stimulus data. The epoch limits differed according to response type as follows: Auditory LLR: -50 to +500 ms; VEP Onset Response: -100 to +500 ms; SSVEP: -500 to +2500 ms. In general, the baseline length was selected to be proportional to the post-stimulus epoch length. For the VEP and SSVEP, the baseline length corresponded to 20% of the post-stimulus epoch time. A relatively short pre-stimulus baseline (50 ms) was used for the Auditory LLR due to the continuous nature of the Cheech, to minimize contamination of previous auditory responses on the current epoch's baseline. Both the Auditory LLR and VEP Onset Response epochs included 500 ms of post-stimulus data, to ensure analysis of all transient ERP components, as well as the sustained negativity observed in the LLR. The SSVEP epoch included data samples for the entire duration of the flickering visual stimulus, which lasted 2.5 seconds. All visual response analyses were time-locked to the start of each 2.5-second flickering stimulus, and the LLRs were time-locked to the voicing onsets in the Cheech, specifically the first chirp in each voiced period.

*Voltage Threshold Artifact Rejection.* Next voltage threshold artifact rejection was done, based on the whole epoch length in all channels except for T7/8, earlobe, and mastoid channels; this excluded any epochs with deflections exceeding ±80 μV from further analysis. This thresholding procedure resulted in the following across-subjects mean percent and mean number of accepted epochs and the across-subjects range of number of accepted epochs, for each response type and visual condition: Auditory LLR: 97% accepted, 1369 mean epochs,

1185-1422 epochs; Visual Onset: 98%, 174 epochs, 152-180 epochs; SSVEP: 96%, 171 epochs, 139-180 epochs.

MLRs and ASSRs.

Following port code correction/addition, noisy segments within the continuous data were removed, corresponding to the same latencies as those excluded for the auditory LLRs and visual response analyses. Next, the data were resampled to 1024 Hz, referenced to the average earlobes, and filtered using a zero-phase, band-pass (0.5 to 200 Hz), $8^{th}$ order Butterworth filter (DC offset was removed prior to filtering). One subject's data were also notch-filtered to remove 60 Hz noise, as previously described. The Independent Component weight matrix calculated for the visual response/auditory LLR analysis stream was applied to the current analysis as well, and the same eye blink component(s) that were removed for the visual/LLR data were also removed from each subject's MLR/ASSR dataset. Next, any bad channels were spatially interpolated (same channels as for the visual/LLR data), and the data were high-pass filtered with a cut-off frequency of 15 Hz using a zero-phase, $8^{th}$ order Butterworth filter, resulting in a passband of 15 to 200 Hz for the MLR/ASSR data. At this point, the auditory port codes were shifted in time to account for the time it takes for sound to travel 1.27 m, from the speaker to the participant (~ 3.7 ms). Next, ERPLAB was used to extract information about the voicing onsets, to which the MLRs and ASSRs were time-locked, and the data were epoched and baselined to the pre-stimulus data. The epoch time limits were as follows: MLR: -5 to +60 ms; ASSR: -150 to +1100 ms. For the MLR, these epoch time limits were originally chosen to encapsulate the entire MLR; recall that there was a 48.8-ms gap between the voicing onset/chirp to which the MLR was time-locked and the next chirp. Like the LLR and ABR, a relatively short baseline period (5 ms) for the MLR was chosen to minimize contamination from residual neural activity due to the continuous auditory stimuli. The ASSR baseline length (150 ms) was selected as a compromise between the long duration of the post-

stimulus epoch length and the need to minimize contamination of the baseline due to the continuous Cheech stimuli. Finally, voltage threshold artifact rejection was done, as previously described, excluding any epochs with deflections exceeding ±80 μV from the ERP averages. This resulted in the following mean percent and mean number of accepted epochs, and range across subjects: MLR: 97% of epochs accepted on average, 1369 mean epochs, 1180-1422 epochs; ASSR: 92% of epochs accepted, 1294 mean epochs, 967-1419 epochs.

ABRs.

For the ABRs, following port code addition/correction, the EEG data files for each block were then concatenated into one file. Noisy data segments, corresponding to the same latencies as those excluded for the other visual and auditory data, were removed. The data were referenced to the average earlobes, and any bad channels were spatially interpolated with a spline function. These bad channels were the same as those identified in the other auditory and visual data. Next, the data were filtered, using a band-pass (100 to 1500 Hz) Butterworth filter (order of 8), and the chirp port codes were shifted in time to account for sound travel time from the speaker to the participant. ERPLAB toolbox was then used to obtain chirp onsets, epoch the data to time-lock to them (epoch limits: -2 to 24 ms) and to baseline the data to the pre-stimulus period. The epoch limits were chosen because there was a minimum of ~24 ms between chirps, and a brief pre-stimulus baseline (2 ms) was used to minimize contamination from residual brainstem activity due to the continuous Cheech stimuli. Threshold artifact detection was conducted to identify and exclude epochs in which activity exceeded ±35 μV, in a subset of channels. ICA was not done for the ABR ERPs, since the ABR passband of 100 to 1500 Hz removes most, if not all of the eye blink artifact. Because the ABR signal is small in comparison to muscle activity and since the ABR peaks at the vertex (Cz), channels near the forehead and temples (which tend to have the most muscle activity) including FP1/2, AF3/4, F7/8, as well as any bad channels and the earlobe and mastoid channels, were excluded from

threshold artifact detection; this was done to preserve as many epochs as possible, with clean EEG signals in the central channels of interest, for creating the ABRs. This resulted in the preservation of an average of 90% of epochs (mean = 10428 epochs, range = 6967-11640 epochs) across subjects.

***Statistical Analyses.***

Custom MATLAB code was used for statistical analysis of the EEG data. Since the goal of this study was to validate the EEG paradigm at the single-subjects level, statistics were run on each individual subject's data, using a bootstrapping approach based on Zhu et al. (2013). This allowed us to quantify the number of subjects that exhibited significant responses to the auditory and visual stimuli. The bootstrapping algorithm differed slightly for different responses, as described below.

ABRs, MLRs, LLRs, and VEPs.

For each subject, the preprocessed, epoched data were imported into MATLAB. For the ABRs and MLRs, the data epochs were shortened to -2 to +15 ms and -5 to +53 ms, respectively. For the ABR, this was done primarily to speed-up computation time of the statistical analysis, since the components of interest occurred within 15 ms. Furthermore, the original MLR epoch included additional time points to +60 ms; however, because the next chirp always occurred at approx. +48.8 ms, the original MLR included the ABR Wave V to the subsequent chirp. Thus, the MLR was truncated to +53 to encompass the Pb and exclude the subsequent Wave V. Pre-stimulus time points were included in the ERP Bootstrapping analysis and are shown in the results figures; this was done for transparency and to ensure that no robust responses were observed due to the continuous nature of especially the auditory stimulus. To generate an estimate of the actual data, a subset of epochs was randomly selected with replacement and the average of this data subset was computed, resulting in an ERP in

each of the 32 scalp channels (excluding earlobe and mastoid channels). This was repeated 100 times, and the grand average of these 100 draws was computed.

To create the null distribution, a subset of actual data epochs was randomly selected with replacement and the amplitude values comprising each epoch were randomly scrambled in time. The mean of these scrambled data epochs was computed. These steps (i.e., draw, scramble, average) were repeated 100 times, and the grand average of these 100 draws was computed. This full process was iterated 1,000 times to generate the null distribution. Since creating the null distribution is computationally expensive, we limited the creation of the null ERPs to only one or two channels as follows: Cz for ABRs, Fz and Cz for MLRs and LLRs, and Oz for VEPs. These channels were chosen based on *a priori* knowledge of the scalp regions where auditory and visual evoked responses generally reach their peak amplitudes (Luck and Kappenman, 2011). The null ERPs were then filtered, using the same filter parameters as done on the actual data, and then baselined to the pre-stimulus period. Each subject's data was used to generate their own null ERPs.

Each individual's null distribution was used to statistically test their own actual data. To control for multiple comparisons across channels and time points, the maximal absolute null value across channels and time points was recorded, resulting in a vector with 1,000 maximal null values. This vector was then sorted in descending order. Next, the absolute value of each actual data point was compared to the sorted maximal null vector, to determine the proportion of null samples that were larger than the absolute value of the actual data point (i.e., resulting in its p-value). This was repeated for each subject and response. Because the maximal null vector comprised 1,000 samples, the minimum p-value possible was 0.001, which was used as the threshold for the single-subject results.

To determine the number of participants with significant responses for each ERP component, we first plotted the group average ERPs and found the peak latency for each observed positive and negative deflection in Channel Cz for ABR, Fz for MLR and LLR, and channel Oz for VEP. Next, using custom MATLAB code, an automated procedure scanned individual subjects' data to ascertain the number of data samples that reached a p-value of 0.001 within a window around the group mean peak for that particular component. The window was defined as the group-mean latency ± 1, 3, or 20 ms for the ABR, MLR, or LLR/VEP responses, respectively. These window durations were chosen in accordance with *a priori* knowledge of the duration of each component peak and confirmed via inspection of the group-averaged ERPs in the present study. Thus, we selected the window durations to account for the increase in peak width from the ABR to MLR to LLR. Since both the VEP and LLR are cortical responses with relatively broad peaks (compared to ABRs and MLRs), we chose a window size of ± 20 ms for consistency in the analysis of both types of cortical responses. For sustained cortical responses (i.e., LLR sustained negativity and VEP late negativity), the window was defined according to the duration of the group-average sustained response. A single-subject significant response was defined as follows: the number of data samples that deflected in the correct direction (positive or negative) and reached a p-value of 0.001 had to exceed one-third of the number of total data samples in the specified window.

SSVEPs and ASSRs.

First, an estimate of the real data was obtained by drawing randomly with replacement, a subset of data epochs. Next, each selected epoch was converted to the frequency domain via a Fast Fourier Transform (FFT) applied from 50 to 1050 ms of the ASSR epochs (i.e., 1 Hz resolution) and from 500 to 2500 ms of the SSVEP epochs (i.e., 0.5 Hz resolution, since the inner ring flickered at 7.5 Hz). For the ASSR, we extracted the data samples starting at 50 ms to avoid the transient portion of the onset response, which arises primarily from subcortical

structures and comprises broadband spectra that overlaps with the ASSR. Likewise, for the SSVEP, we used the data samples starting at 500 ms to avoid the visual onset response, whose spectral energy overlaps with that of the flicker rates. The single-sided FFT was computed for each epoch and scaled to the number of data samples on which the FFT was performed. The mean single-sided FFT (complex-values) was computed across the subset of data epochs drawn, and the absolute value was taken to obtain magnitude. These magnitude values were then converted to dB (arbitrary units). This procedure was repeated 100 times (i.e., number of draws), and the grand average magnitude was calculated across these 100 draws to obtain the estimated magnitude of the actual data in each of the 32 scalp electrodes. This was repeated for each participant's data.

To create the null distribution for each participant, the same steps were followed for the actual data. However, after computing the FFT for each epoch, the magnitude component of the FFT was preserved, but the phase was randomized from 0 to 2*pi. In theory, this should provide an accurate estimate of the noise floor in the data (Zhu et al. 2013). Thus, phase-randomized FFTs were obtained, using the actual data's magnitude component and the random phase vector. Next, just like the actual data, the phase-randomized FFTs were converted to the single-sided spectra, scaled, and averaged across epochs. The magnitude component was extracted and converted to dB (arbitrary units). The grand average dB magnitude was computed across 100 draws, and this whole procedure was repeated 1,000 times to create the null (phase-randomized) distribution. Due to the computational cost of creating the null distribution, this was limited to channels Fz and Cz for the ASSRs and to channel Oz for the SSVEPs. Like the ERP analysis, these channels were chosen based on *a priori* knowledge of the scalp regions where auditory and visual evoked potentials generally reach their peak amplitudes (Luck and Kappenman, 2011). For data visualization, the signal-to-noise ratio

(SNR) was obtained by subtracting the mean of the null distribution (i.e., the noise floor, in dB units) from the mean of the estimated actual data (in dB units), for each participant.

Statistics were performed at the single-subjects level. For each channel (Fz and Cz, or Oz) and frequency of interest (i.e., ASSR: 41 Hz ($f_0$), 82 Hz, 123 Hz, 164 Hz; SSVEP: 7.5 Hz (Inner ring $f_0$), 12 Hz (Outer ring $f_0$), 15 Hz (Inner ring harmonic), 24 Hz (Outer ring harmonic)), the actual data value was compared to the distribution of null values at that frequency to determine its p-value. A Bonferroni-corrected threshold was computed as 0.05/(f*c), where f is the number of frequencies of interest (i.e., 4) and c is the number of channels examined (i.e., 2 for ASSR, 1 for SSVEP), resulting in thresholds of 0.00625 for ASSR and 0.0125 for SSVEP. These Bonferroni-corrected thresholds were used to determine the number of subjects with a significant response at each frequency of interest.

Number of Epochs for Bootstrapping.

The number of epochs drawn for each response type was initially based on the minimum number of epochs available after artifact rejection across subjects, so that all subjects could be included in all analyses. These epoch numbers were further reduced, to accommodate the minimum number of artifact-free trials that we expect (and have obtained) from young children participating in the same EEG paradigm. For all visual responses (VEP, SSVEP), 50 epochs were drawn (randomly, with replacement) for each iteration of the bootstrapping analysis. For all auditory responses (ABR, MLR, LLR, and ASSR), 500 epochs were selected (randomly, with replacement).

Assessing Relationships among Auditory and Visual EEG Responses.

As a supplementary analysis, we conducted across-subjects correlations to determine if the various auditory and visual EEG responses varied in a systematic way. To do this, we converted the amplitude estimates of a subject's true data for each EEG response into z-scores,

627 relative to the mean and standard deviation of each subject's null distribution. For each ERP
628 response, each subject's null distribution mean and standard deviation were computed across
629 all time points and channels for which the null was computed. For the steady-state responses,
630 each subject's null distribution mean and standard deviation were computed using the one or
631 two channels for which the null was created, but separately for each frequency of interest. By
632 converting the data to z-scores, we could directly compare different responses with different
633 magnitudes or measurement units (see also Zhu et al. 2013). For the ERP responses (ABR,
634 MLR, LLR, VEP), each individual's peak z-score was obtained for each component, using the
635 same windowing procedure as described for quantifying the number of subjects with a
636 significant ERP response. Data from one channel were used from each EEG response: Cz for
637 ABR, Fz for MLR, LLR, and ASSR, and Oz for VEP and SSVEP. For any observed sustained
638 potentials, each individual's z-scores across the time range of the group-mean sustained
639 potential were averaged. Next, the negative components' z-scores were multiplied by -1, and
640 the z-scores corresponding to each component within a given response were averaged (e.g.,
641 P1, N1, and sustained negativity for LLR). This was done to create an aggregate z-score for
642 each EEG response. Similarly, for the ASSR and SSVEP, the z-scores corresponding to each
643 frequency of interest were averaged.

644 Using MATLAB, Pearson correlations (two-tailed) were conducted across-subjects for
645 each possible pair of auditory responses (ABR-MLR, ABR-LLR, ABR-ASSR, MLR-LLR, MLR-
646 ASSR, and LLR-ASSR) and between the visual responses (VEP-SSVEP). Also, to assess if the
647 frequencies of interest were correlated within the ASSR and SSVEP responses, pairwise
648 Pearson correlations were computed on the z-scores of each frequency of interest (ASSR: 41-
649 82 Hz, 41-123 Hz, 41-164 Hz, 82-123 Hz, 82-164 Hz, 123-164 Hz; SSVEP: 7.5-12 Hz, 7.5-15
650 Hz, 7.5-24 Hz, 12-15 Hz, 12-24 Hz, 15-24 Hz).

651

## Results

### ERP Responses (ABR, MLR, LLR, VEP)

**Table 1** contains a summary of the number of subjects showing a significant response for each observed component, along with peak amplitude and latency measurement results.

Examination of the ABR data revealed two negative peaks interleaved with two positive deflections (**Figure 2**). The first negative peak, which we have labeled "$n_0$", occurred around 3.5 ms and was maximal over fronto-central channels (significant for 24 subjects). Next, we observed a positive peak at 6.5 ms, which was maximal over the vertex, consistent with the timing and topography of Wave V (significant for all 25 subjects). Another negative peak followed around 9 ms (significant for all 25 subjects), and a positive peak at 13 ms, which were maximal over frontal sites, suggesting a neural generator in/near auditory cortex. These peaks' timings are consistent with the $N_0$ and $P_0$ components, respectively (Mendel and Goldstein 1969; Picton et al. 1974; Yoshiura et al. 1996), indicating the transition between the ABR and MLR. Of the four deflections observed, the $P_0$ was by far the weakest in terms of amplitude and number of subjects with a significant response (i.e., 16 subjects; see **Table 1**).

Analysis of the MLR data (**Figure 3**) showed that at the group level, all MLR components were evident (Na, Pa, Nb, Pb), along with ABR Wave V, which had a slightly later, broader peak than in the ABR analysis. This shift in latency is likely due to the different bandpass filters applied to the ABRs (100-1500 Hz) and MLRs (15-200 Hz), such that low-frequency activity dominates the MLR representation of ABR Wave V. As shown in **Figure 3B**, the MLR components peaked over frontal sites. At the single-subjects level in channel Fz, the majority of participants had significant MLR components, but the Na and Pb were the most robust in terms of amplitude and number of subjects with a significant response (22 and 20 subjects, respectively). In channel Cz, there were 17, 15, 11, and 19 subjects showing significant Na, Pa,

676 Nb, and Pb responses, respectively. Furthermore, examination of the single-subject data in Fz
677 (p threshold of 0.001) revealed that all 25 subjects had at least one significant MLR component,
678 24 had at least two significant components, and 22 had at least three significant MLR
679 components.

680 The LLR data revealed a P1 that peaked at ~80 ms and was relatively broad in latency,
681 followed by an N1 that peaked at ~170 ms and a sustained negativity that was evident from
682 about 225 to 425 ms after voicing onsets in the Cheech (**Figure 4**). All three components had
683 fronto-central topography, suggestive of auditory cortex neural generators. Currently, it is
684 unclear if the observed sustained negativity reflects truly sustained activity and/or overlapping
685 N1's due to the continuous nature of the auditory stimulus. However, the topography of the
686 sustained negativity is very similar to the N1 topography. In channel Fz, all 25 subjects had a
687 significant P1, 21 had a significant N1, and 23 had a significant sustained negativity. In channel
688 Cz, the results were similar, but slightly weaker; 25, 19, and 22 subjects had a significant P1,
689 N1, and sustained negativity, respectively.

690 As illustrated in **Figure 5**, analysis of the VEP data revealed the visual P1, N1, and P2
691 components, followed by a late negativity over posterior sites, which started around 420 ms and
692 continued to the end of the epoch period. A subsequent examination of the group data, which
693 used a longer epoch (to 3 seconds beyond flicker onset), showed that this posterior negativity
694 continued until 770 ms after flicker onset. Inspection of the group-average scalp topographies
695 revealed that the visual P1 and N1 peaked over posterior sites, whereas the P2 showed a broad
696 scalp distribution that was maximal over midline sites across the scalp. At the single-subjects
697 level, a majority of participants had significant responses for each of the four components
698 identified. The P1 was the strongest in terms of number of subjects showing a significant
699 response (22 subjects), followed by the P2 and sustained negativity (20 subjects), and finally
700 the N1 (17 subjects). Furthermore, examination of the results in channel Oz (p-threshold of

0.001) revealed that all 25 subjects showed at least two VEP components significantly, and 19 had at least three components.

**Steady-State Responses (ASSR, SSVEP)**

**Table 2** contains a summary of the number of subjects showing a significant response for each frequency of interest, along with the group raw magnitude and signal-to-noise ratio (SNR) results.

Inspection of the ASSR data revealed large peaks at the stimulation frequency (41 Hz) and its three harmonics (82 Hz, 123 Hz, 164 Hz), as displayed in **Figure 6**. At the group level, the scalp location of maximum amplitude differed among the ASSR frequencies, with the lowest frequency (41Hz) peaking fronto-centrally and the highest frequency (164Hz) peaking at the vertex, possibly reflecting differential contributions from the ascending auditory pathway, in line with Herdman et al. (2002) and Coffey et al. (2016). In the present study, the ASSR (and SSVEP) raw magnitudes were first estimated by converting each epoch's time waveforms to the frequency domain via a Fourier transform and then averaging across the frequency spectra, within a bootstrapping algorithm. Similarly, the noise floor was modeled in similar fashion, with the added step of randomizing phase (but preserving magnitude) before averaging frequency spectra, as described in the Methods section. These raw magnitude and noise floor data were used for statistical thresholding at the four frequencies of interest in channels Fz and Cz and were used to compile the data in **Table 2**. All 25 subjects had significant ASSRs at 41, 123 and 164 Hz, and 24 of 25 subjects had significant responses at 82 Hz; this pattern of results was observed in both Fz and Cz.

As shown in Figure 6B, the noise floor estimate accurately modeled the 1/f shape of the noise floor in the actual ASSR data. However, the noise floor estimate was uniformly lower than the raw magnitude estimate (red vs. blue line in Figure 6B), resulting in the floor of the signal-to-

noise ratio (SNR) hovering around 1.75 dB (black line) instead of 0 dB. This is suggestive of broad-spectrum, weakly phase-locked neural activity that is driving the noise floor of the actual ASSR data above what would be expected by chance (i.e., random phase across epochs). To further probe this issue, we used a similar bootstrapping procedure to estimate the actual ASSR data, but the epochs were averaged in time for each draw, and subsequently across 100 draws, before converting it to frequency space. By averaging in time first, any weakly phase-locked activity should be attenuated due to destructive interference. The resulting group average raw magnitude spectrum is depicted in Figure 6B (gray line). Indeed, its noise floor is much lower than both the noise floor estimate and original raw magnitude spectrum, while the peaks of both raw magnitude spectra reached nearly identical values in channel Fz. This corroborates the notion that weakly phase-locked activity contributes to the raw magnitude spectrum (and noise floor estimation) averaged in the frequency domain.

Examination of the SSVEP data revealed peaks at the stimulation rates (7.5, 12 Hz) and their first harmonics (15, 24 Hz), that were maximal over parieto-occipital sites (**Figure 7**). At the individual subjects level, 21, 24, 22, and 19 subjects had significant neural responses at 7.5, 12, 15, and 24 Hz, respectively, in channel Oz as shown in **Table 2**. Furthermore, 20 participants had significant neural responses at both 7.5 and 12 Hz (stimulation fundamental frequencies), and the other 5 subjects had significant responses at either 7.5 or 12 Hz. However, like the ASSR, the noise floor estimate accurately reflected the 1/f shape of the noise floor of the SSVEP magnitude spectrum, but it was uniformly ~1.2 dB lower than the apparent noise floor in the actual SSVEP magnitude spectrum, obtained by averaging in frequency space. Thus, we also computed SSVEP magnitude spectra, by averaging data in the time domain first before converting to the frequency domain (Figure 7B, gray line). Like the ASSR, the apparent noise floor of the time-averaged SSVEP spectrum dropped below that of the estimated noise floor, again suggesting the contribution of weakly phase-locked neural activity

to the raw magnitude spectrum and noise floor estimation, which were averaged in the frequency domain.

**Summary of ERP and Steady-State Response Results**

In summary, the paradigm was successful in eliciting multiple auditory and visual responses across the subjects tested. For the ABR, LLR, ASSR, and SSVEP there was strong convergence across subjects regarding which components were reliably detected. For the MLR and VEP, the pattern was more heterogeneous; while nearly all subjects showed at least two or three of the components within the MLR or VEP, these components were not necessarily the same across participants. Nevertheless, these results demonstrate the robustness of our novel paradigm.

**Assessing Relationships among Auditory and Visual EEG Responses**

We created individual aggregate z-scores for each of the auditory and visual responses, by finding the z-score at the single-subject component peaks or for sustained responses, averaging the z-scores across a pre-defined time range (i.e., 227-426 ms for LLR sustained negativity; 422-496 for VEP late negativity). All ERP components described in the previous section were included in these aggregate z-scores as follows: ABR: $n_0$, Wave V, $N_0$, $P_0$; MLR: Na, Pa, Nb, Pb; LLR: P1, N1, sustained negativity; VEP: P1, N1, P2, late negativity. For the ASSR and SSVEP, the individual z-scores were averaged across the four frequencies of interest to create aggregate z-scores. All aggregate z-scores for each participant are shown in **Figure 8**.

First, we conducted across-subjects pairwise Pearson correlations across the four auditory responses (ABR, MLR, LLR, ASSR). Next, an across-subjects correlation was computed between the VEP and SSVEP z-scores. Finally, we performed correlations within the ASSR and SSVEP z-scores (all 25 subjects included), to determine if the magnitude at the

774 frequencies of interest were systematically related. None of the correlations were significant
775 after controlling for multiple comparisons, except for the correlation between the MLR and
776 ASSR ($r = 0.61$, $p = 0.001$, uncorrected) – which was likely driven by one participant. Therefore,
777 we did not observe any reliable systematic relationships among the different auditory and visual
778 EEG responses.

779 **Discussion**

780 We have developed a novel EEG paradigm to simultaneously record neural activity from
781 visual cortex, and from both subcortical and cortical auditory structures using a continuous
782 speech stimulus, in an unprecedentedly brief amount of time (6-12 min). To determine the
783 efficacy of this new paradigm, we have reported data from a group of healthy young adults.
784 Overall, most participants had significant responses for each of the components examined,
785 despite the conservative null distributions and the stringent p-value thresholds used for
786 statistical testing. The ABR, LLR, and ASSR tended to be the most robust responses, such that
787 all 25 participants had a significant response for the following components: ABR Wave V and
788 $N_0$, Auditory P1, and ASSR (41, 123, and 164 Hz). In terms of the number of participants who
789 showed significant responses, the SSVEP was next, followed by the VEP, and finally the MLR.
790 Furthermore, as described in the Results section, the pattern of results was most
791 heterogeneous for the VEP and MLR. While all 25 participants showed a significant response
792 for at least one MLR component and at least two VEP components, these were not always the
793 same MLR or VEP components across participants. This heterogeneous pattern partly reflects
794 our stringent statistical criteria and highlights the importance of examining the data, especially
795 the MLR and VEP, at the single-subjects level, particularly if this paradigm is used with clinical
796 populations. It also points to potential attributes of the audio and video stimuli that may be
797 further optimized to yield even more consistent responses across all subjects.

Based on an inspection of the ERP waveforms at the group level, the present paradigm generally elicited the canonical responses, in terms of waveform morphology. The only exceptions were the ABR, for which we did not observe the early waves before Wave V, and the non-canonical LLR (which was expected). The ABR result is in contrast to a previous implementation of the Cheech approach by our group, in which the early waves were observed, using monaural delivery of the auditory stimuli via an insert earphone (Miller et al. 2017). Thus, the lack of early waves in the present implementation may reflect the fact that the Cheech was presented in the free-field. Furthermore, for the LLR, we observed the P1, N1, and a sustained negativity, but not the P2 component, in response to voicing onsets within the continuous Cheech. Notably, this morphology has been previously observed in studies using long-duration sounds (Krishnan et al. 2012; Picton et al. 1978a; b), suggesting that this morphology is typical when employing long-duration or continuous auditory stimuli. The present paradigm also successfully evoked neural activity at the auditory stimulation rate (41 Hz) and its first three harmonics (82, 123, and 164 Hz), in addition to neural activity at the visual stimulation rates (7.5 and 12 Hz) and their first harmonic (15 and 24 Hz). Taken together, these results demonstrate that it is feasible to obtain all responses simultaneously, despite stimulating both auditory and visual systems concurrently.

One advantage of obtaining a variety of responses simultaneously within individual participants is that it allows for the assessment of relationships among the various responses. Thus, in the present study, we conducted a series of correlations to understand if and how different EEG responses' amplitude (converted to z-scores) related to one another. Overall, we did not observe any robust across-subjects correlations among the different EEG responses. This generally suggests that the examined responses reflect different processing operations and/or different neural generators. Furthermore, assuming some inter-subject variability in EEG recording quality, the lack of uniform relationships across responses indicates that recording

quality variability is unlikely to induce false across-subjects correlations among the different EEG responses. Moreover, by examining all responses simultaneously, the lack of systematic relationships among the EEG responses within subjects cannot be due to changes in for example, brain state or alertness, across time, as is the case for serial recording paradigms. Taken together, these points highlight the importance of assessing all responses simultaneously for a thorough evaluation of the auditory and visual systems – which the current EEG paradigm enables.

Previously, a variety of paradigms have been developed to record EEG activity at multiple processing levels, using auditory stimuli. Here, we compare these previous paradigms to the current one, in terms of the responses recorded, whether the different responses were recorded simultaneously or serially, the type of stimuli used, and the EEG recording duration.

First, regarding the responses recorded, previous paradigms have mostly recorded the subcortical (usually frequency-following response (FFR)) and cortical (i.e., LLR) activity (Bidelman 2015; Bidelman and Alain 2015; Bidelman et al. 2013; Bidelman et al. 2014a; Bidelman et al. 2014b; Krishnan et al. 2012; Musacchia et al. 2008). Woods et al. (1993) analyzed the ABR, MLR, and LLR; similarly, Shiga et al. (2015) developed a paradigm to examine the FFR, MLR, and LLR (MMN). Slugocki and colleagues (2017) measured a variety of subcortical and cortical responses, including FFR, 40- and 80-Hz ASSR, LLR, MMN, and P3a. Finally, Sculthorpe-Petley et al.'s (2015) paradigm measured only cortical responses, including LLR (N1), MMN, P300, N400, and Early Negative Enhancement (reflects recognition of hearing one own's name; Holler et al. 2011; Tateuchi et al. 2012). In contrast, the present paradigm enables the recording of subcortical and cortical auditory activity (ABR, MLR, LLR, ASSR), as well as cortical visual activity (VEP, SSVEP).

With respect to how both auditory subcortical and cortical responses were recorded, various approaches have been used. Subcortical responses occur earlier and thus necessitate smaller inter-stimulus intervals (ISIs) and higher EEG acquisition sampling rates than cortical responses. Many studies have recorded brainstem and cortical responses sequentially in separate blocks (e.g., Bidelman and Alain 2015; Bidelman et al. 2013; Bidelman et al. 2014a; Bidelman et al. 2014b; Musacchia et al. 2008) or in interleaved clusters (Bidelman 2015); these approaches usually involve using shorter ISIs for the brainstem blocks/clusters than the cortical blocks/clusters. Other studies have recorded auditory brainstem and cortical responses simultaneously, using fixed ISIs (e.g., Krishnan et al. 2012; Shiga et al. 2015; Slugocki et al. 2017) or variable ISIs (e.g., 40 to 200 ms; Woods et al. 1993) to accommodate both types of responses.

Regarding the types of auditory stimuli used, previous studies have employed amplitude-modulated tones (Shiga et al. 2015; Slugocki et al. 2017), tone pips in the midst of broadband masking noise (Woods et al. 1993), iterated rippled noise stimuli (Krishnan et al. 2012), and synthetic vowel or consonant-vowel stimuli (e.g., Bidelman 2015; Bidelman et al. 2013; Musacchia et al. 2008). Sculthorpe-Petley and colleagues' (2015) paradigm used tones in one half of the recording, and continuous speech (sentences) in the other half. In the present study, the use of chirps embedded into continuous speech (Cheech) allows for the simultaneous recording of subcortical and cortical activity in response to a naturalistic stimulus.

With respect to EEG recording duration, the fastest of these studies was that by Petley-Sculthorpe et al. (2015), which approximated only 5 minutes; however, only cortical responses were recorded. For paradigms involving recording both subcortical and cortical responses, Bidelman's (2015) clustering approach took approximately 28 minutes, while the paradigm described in Shiga et al. (2015) lasted around 38 minutes. Likewise, Slugocki and colleagues' (2017) paradigm involved about 40 minutes of recording time. In contrast, the paradigm

described herein involves 12 minutes maximum of recording time – which is considerably faster than these other approaches devised to collect both subcortical and cortical auditory responses.

In fact, the present data were collected in only 12 minutes, mainly to allow for enough trials for the originally planned black-white versus red-green comparison (6 minutes per color condition). We have used a 10-minute black-white-only version of this paradigm to collect data in young children, which is ample time to yield reliable ERPs for most children, even after noisy data segments and epochs were removed (unpublished data from our laboratory). The brief time required makes this paradigm ideal for individuals who are unable to sit through a long study (e.g., toddlers), and allows for short study sessions, which is advantageous to both busy participants and researchers.

Furthermore, in the present EEG paradigm, participants watched cartoon clips during the presentation of the visual flicker and auditory Cheech stimuli. This was done to render the paradigm infant/child friendly. Also, because this task requires no behavioral responses, it can be used in infants and young children, as well as in individuals with limited communication abilities. That said, the paradigm can also easily be adapted into an active task, for instance to investigate top-down attention effects on the various auditory and visual responses recorded.

Along these lines, one caveat of the present implementation is that although participants were instructed to watch the cartoon clips, they were not explicitly told to ignore the flicker or Cheech stimuli. Thus, it is important to acknowledge that participants' attention likely wandered to the flicker or Cheech stimuli at times, and consequently, it is important to acknowledge the known effects of attention on the measured responses. Selective attention to or away from a particular auditory stimulus has been shown to have little to no effect on the ABR (Hackley et al. 1990; Woldorff and Hillyard 1991), a small effect on the MLR, particularly after 20 ms (Hackley et al. 1990; Woldorff et al. 1987), and the most robust effect on the LLR, particularly the N1 and

P2 components (Hillyard et al. 1973; Picton et al. 1971). Similarly, attending (or not) to a visual stimulus affects the VEP, especially from the P1 and later (e.g., Clark and Hillyard 1996; Gomez Gonzalez et al. 1994; Mangun et al. 1993). Furthermore, there is evidence that selective attention can enhance the 40-Hz ASSR to the attended sound's stimulation rate (Bharadwaj et al. 2014; Tiitinen et al. 1993) (but see Mahajan et al. 2014; Muller et al. 2009) and the SSVEP to the attended visual stimulus' stimulation rate (Andersen et al. 2015; Morgan et al. 1996). Thus, these findings indicate that selective attention has a stronger effect on the neural response to a stimulus as it ascends from sub-cortical to cortical processing regions.

Since attention was not explicitly manipulated in the current paradigm, future studies can use this paradigm in conjunction with an attention manipulation, to quantify how selective attention modulates the recorded responses. Furthermore, in the present study, participants watched a cartoon while also perceiving flickering visual stimuli. Thus, it is possible that if participants generally devoted more visual attention to the cartoon than the visual flickers, then they would elicit smaller VEPs and SSVEPs than if these stimuli had been actively attended. This explanation may account for the fact that the VEP and SSVEP measures were significantly detected in fewer subjects overall than the ABR, LLR, and ASSR. Despite this caveat, we observed the VEP components and SSVEP responses in a majority of participants; this suggests that even if reduced attention to the flickers has an effect on the visual responses, it does not eradicate them in the present paradigm.

Furthermore, this paradigm is flexible, in that the visual and auditory stimuli can be customized according to one's particular research questions. For example, including audio-only and visual-only periods of stimulation, along with concurrent auditory and visual stimulation, one could directly examine how activity may be modulated by the stimulus context in an individual. Moreover, the concurrent auditory and visual stimuli also present opportunities for the examination of complex interactions between the auditory and visual systems. For example,

this use of multisensory stimulation may provide new and important data on one's ability to deal with multiple and competing sensory information. However, the multisensory nature of this paradigm is not limited to competing or distracting visual stimuli. Alternatively, future implementations of this paradigm could use congruent and incongruent talking face stimuli that coincide with the Cheech, to study neural mechanisms involved in audiovisual integration of speech. Regarding feasibility in terms of accurate EEG trigger timing, careful attention to timing is important when implementing the current EEG paradigm. First, when implementing the present paradigm, it is important to quantify timing differences between the trigger codes and stimuli and to determine if any auditory trigger codes are being missed, using an oscilloscope. Next, if timing inconsistencies are observed, there are various ways to address this. One way would be to adjust triggers and add missing triggers post-hoc, using the approach that we have developed; our MATLAB code for this approach has been posted on GitHub at https://github.com/MillerLab-UCDavis/Cheech-Toolbox. There are also other ways to obtain accurate trigger timing; for example, one could play a third audio channel that contains only trigger pulses and send that signal to the EEG system directly, or one could use a third-party device designed for accurate trigger timing, such as the Cedrus StimTracker. Therefore, obtaining precise trigger timing when implementing the present EEG paradigm is feasible.

There are various potential applications of this paradigm to clinical research, even beyond assessing auditory function in the context of hearing loss. For example, its utility may be investigated in patients with multiple sclerosis (MS), as previous studies have indicated usefulness of auditory evoked potentials (especially ABRs and MLRs) for detecting neurological abnormalities in some cases (e.g., Celebisoy et al. 1996; Japaridze et al. 2002; Soustiel et al. 1996). Moreover, this paradigm would additionally allow for the assessment of visual cortical responses in these patients, especially since visual disturbances are one of the most common reported signs in patients with MS (Milner et al. 1974). In another context, the multisensory

nature of the paradigm can be harnessed to investigate interactions between auditory and visual processing in children with hearing loss (Backer et al. 2017) or in children with autism spectrum disorder or auditory and language processing disorders. These are just a few examples of many potential applications.

Future research involving this paradigm, especially that involving clinical populations, will provide valuable information regarding the generalizability of the present results. We should make clear that the technique is free and unrestricted for non-commercial research and educational use, in accordance with the University of California's "Principles Regarding Rights to Future Research Results" guidelines (see Principle #3 at https://www.ucop.edu/research-policy-analysis-coordination/_files/Principles%20Guidelines.pdf). For more information or to inquire about a license for commercial use or commercial applications, please contact the UC Davis Office of Research at innovation*Access*@ucdavis.edu.

In conclusion, we have demonstrated the use of a new EEG paradigm to concurrently stimulate and record subcortical and cortical auditory activity, as well as parafoveal and peripheral cortical visual activity, in about 6-12 minutes. In light of the short recording time and the flexibility to customize the auditory and visual stimuli depending on the study's population and goals, this EEG paradigm may be useful for both basic and clinical research objectives.

## Acknowledgements

The authors are grateful to Todd LaMarr, Ludmila Ciochina, Surah Alsawaf, Danielle Fujino, and Alaina Porter for their assistance with data collection. Thanks to the House Ear Institute for providing the Harvard/IEEE recordings.

Please note Laurel A. Lawyer's present affiliation: Department of Language and Linguistics, University of Essex, Colchester, UK. Andrew S. Kessler's present affiliation: Verathon Inc., Division of Imaging and Scanning Solutions, Bothell, WA.

## Grants

Funding from the National Institutes of Health, National Institute on Deafness and Other Communication Disorders (Grant DC014767 awarded to DPC), the Child Family Fund for the Center for Mind and Brain (awarded to LMM and DPC), and the UC Davis Science Translation and Innovative Research (STAIR) grant (awarded to LMM) supported this research.

## Disclosures

As noted in the manuscript, the chirp-speech approach is patent-pending and owned by the Regents of the University of California, with Lee M. Miller as inventor. The University of California encourages the use of this technique for non-commercial educational and research purposes. For more information or to inquire about a license for commercial use or commercial applications, please contact the UC Davis Office of Research at innovation*Access*@ucdavis.edu. No financial or other conflicts of interest (e.g., licensing agreements) exist.

## Footnotes

[1] In many reports, the terms “Pb” (MLR component) and “P1” (LLR component) are used interchangeably, due to overlapping time courses. However, the Pb and P1 may have different neural generators and developmental trajectories (Ponton et al. 2002), suggesting that they may index different aspects of auditory processing.

[2] Acoustic periodicity can also elicit another type of auditory response, the frequency following response (FFR) (for reviews, see Krishnan 2007; Skoe and Kraus 2010), which is traditionally thought to have neural generators in sub-cortical structures (Chandrasekaran and Kraus 2010), but recent studies suggest additional contributions from the auditory cortex (Coffey et al. 2016), if stimulus frequency is relatively low (Bidelman 2018). However, further discussion of the FFR is beyond the scope of the present report.

## References


IEEE Recommended Practice for Speech Quality Measurements. *IEEE No 297-1969* 1-24, 1969.

**Aine CJ, Supek S, and George JS**. Temporal dynamics of visual-evoked neuromagnetic sources: effects of stimulus parameters and selective attention. *Int J Neurosci* 80: 79-104, 1995.

**Andersen SK, Hillyard SA, and Muller MM**. Attention facilitates multiple stimulus features in parallel in human visual cortex. *Curr Biol* 18: 1006-1009, 2008.

**Andersen SK, Muller MM, and Hillyard SA**. Attentional Selection of Feature Conjunctions Is Accomplished by Parallel and Independent Selection of Single Features. *J Neurosci* 35: 9912-9919, 2015.

**Arslan E, Prosser S, and Michelini S**. Simultaneous recording of auditory evoked potentials. Relationships among the fast, middle and long latency components. *Scand Audiol* 13: 75-81, 1984.

**Artieda J, Valencia M, Alegre M, Olaziregi O, Urrestarazu E, and Iriarte J**. Potentials evoked by chirp-modulated tones: a new technique to evaluate oscillatory activity in the auditory pathway. *Clin Neurophysiol* 115: 699-709, 2004.

**Backer KC, Kessler AS, Coffey-Corina S, Lawyer LA, Miller LM, and Corina DP**. Abstract: Charting developmental changes in auditory and visual evoked potentials in children with cochlear implants. In: *Conference on Implantable Auditory Protheses*. Lake Tahoe, CA: 2017.

**Bell AJ, and Sejnowski TJ**. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput* 7: 1129-1159, 1995.

**Bell SL, Allen R, and Lutman ME**. An investigation of the use of band-limited chirp stimuli to obtain the auditory brainstem response. *Int J Audiol* 41: 271-278, 2002a.

**Bell SL, Allen R, and Lutman ME**. Optimizing the acquisition time of the middle latency response using maximum length sequences and chirps. *J Acoust Soc Am* 112: 2065-2073, 2002b.

**Bharadwaj HM, Lee AK, and Shinn-Cunningham BG**. Measuring auditory selective attention using frequency tagging. *Front Integr Neurosci* 8: 6, 2014.

**Bidelman GM**. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage* 175: 56-69, 2018.

**Bidelman GM**. Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials. *J Neurosci Methods* 241: 94-100, 2015.

**Bidelman GM, and Alain C**. Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J Neurosci* 35: 1240-1249, 2015.

**Bidelman GM, Moreno S, and Alain C**. Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage* 79: 201-212, 2013.

**Bidelman GM, Villafuerte JW, Moreno S, and Alain C**. Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol Aging* 35: 2526-2540, 2014a.

**Bidelman GM, Weiss MW, Moreno S, and Alain C**. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur J Neurosci* 40: 2662-2673, 2014b.

**Boersma P, and Weenink D**. *PRAAT, a system for doing phonetics by computer*. 2001, p. 341-345.

**Burkitt GR, Silberstein RB, Cadusch PJ, and Wood AW**. Steady-state visual evoked potentials and travelling waves. *Clin Neurophysiol* 111: 246-258, 2000.

**Celebisoy N, Aydogdu I, Ekmekci O, and Akurekli O**. Middle latency auditory evoked potentials (MLAEPs) in (MS). *Acta Neurol Scand* 93: 318-321, 1996.

**Celesia GG**. Organization of auditory cortical areas in man. *Brain* 99: 403-414, 1976.
**Chandrasekaran B, and Kraus N**. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47: 236-246, 2010.
**Clark VP, Fan S, and Hillyard SA**. Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping* 2: 170-187, 1994.
**Clark VP, and Hillyard SA**. Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *J Cogn Neurosci* 8: 387-402, 1996.
**Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, and Zatorre RJ**. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun* 7: 11070, 2016.
**Cowey A, and Rolls ET**. Human cortical magnification factor and its relation to visual acuity. *Exp Brain Res* 21: 447-454, 1974.
**Daniel PM, and Whitteridge D**. The representation of the visual field on the cerebral cortex in monkeys. *J Physiol* 159: 203-221, 1961.
**Dau T, Wegner O, Mellert V, and Kollmeier B**. Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J Acoust Soc Am* 107: 1530-1540, 2000.
**Davis H, and Zerlin S**. Acoustic relations of the human vertex potential. *J Acoust Soc Am* 39: 109-116, 1966.
**Davis PA**. Effects of acoustic stimuli on the waking human brain. *Journal of Neurophysiology* 2: 494-499, 1939.
**Delorme A, and Makeig S**. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134: 9-21, 2004.
**Di Russo F, Martinez A, Sereno MI, Pitzalis S, and Hillyard SA**. Cortical sources of the early components of the visual evoked potential. *Hum Brain Mapp* 15: 95-111, 2002.
**Di Russo F, Pitzalis S, Spitoni G, Aprile T, Patria F, Spinelli D, and Hillyard SA**. Identification of the neural sources of the pattern-reversal VEP. *Neuroimage* 24: 874-886, 2005.
**Ding J, Sperling G, and Srinivasan R**. Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. *Cereb Cortex* 16: 1016-1029, 2006.
**Elberling C, and Don M**. Auditory brainstem responses to a chirp stimulus designed from derived-band latencies in normal-hearing subjects. *J Acoust Soc Am* 124: 3022-3037, 2008.
**Elberling C, Don M, Cebulla M, and Sturzebecher E**. Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay. *J Acoust Soc Am* 122: 2772-2785, 2007.
**Galambos R, Makeig S, and Talmachoff PJ**. A 40-Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A* 78: 2643-2647, 1981.
**Geisler CD, Frishkopf LS, and Rosenblith WA**. Extracranial responses to acoustic clicks in man. *Science* 128: 1210-1211, 1958.
**Goldstein R, and Rodman LB**. Early components of averaged evoked responses to rapidly repeated auditory stimuli. *J Speech Hear Res* 10: 697-705, 1967.
**Gomez Gonzalez CM, Clark VP, Fan S, Luck SJ, and Hillyard SA**. Sources of attention-sensitive visual event-related potentials. *Brain Topogr* 7: 41-51, 1994.
**Hackley SA, Woldorff M, and Hillyard SA**. Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. *Psychophysiology* 27: 195-208, 1990.
**Halgren E**. Human Evoked Potentials. In: *Neurophysiological Techniques: Applications to Neural Systems*, edited by Boulton AA, Baker GB, and Vanderwolf CH. Totowa, NJ: Humana Press, 1990, p. 147-275.
**Hall JW**. *New Handbook of Auditory Evoked Responses*. Pearson, 2007.
**Hari R, Aittoniemi K, Jarvinen ML, Katila T, and Varpula T**. Auditory evoked transient and sustained magnetic fields of the human brain. Localization of neural generators. *Exp Brain Res* 40: 237-240, 1980.
**Hashimoto I**. Auditory evoked potentials from the human midbrain: slow brain stem responses. *Electroencephalogr Clin Neurophysiol* 53: 652-657, 1982.

**Hashimoto I, Ishiyama Y, and Tozuka G**. Bilaterally recorded brain stem auditory evoked responses. Their asymmetric abnormalities and lesions of the brain stem. *Arch Neurol* 36: 161-167, 1979.
**Herdman AT, Lins O, Van Roon P, Stapells DR, Scherg M, and Picton TW**. Intracerebral sources of human auditory steady-state responses. *Brain Topogr* 15: 69-86, 2002.
**Hillyard SA, Hink RF, Schwent VL, and Picton TW**. Electrical signs of selective attention in the human brain. *Science* 182: 177-180, 1973.
**Holler Y, Kronbichler M, Bergmann J, Crone JS, Ladurner G, and Golaszewski S**. EEG frequency analysis of responses to the own-name stimulus. *Clin Neurophysiol* 122: 99-106, 2011.
**Itthipuripat S, Garcia JO, and Serences JT**. Temporal dynamics of divided spatial attention. *J Neurophysiol* 109: 2364-2373, 2013.
**Japaridze G, Shakarishvili R, and Kevanishvili Z**. Auditory brainstem, middle-latency, and slow cortical responses in multiple sclerosis. *Acta Neurol Scand* 106: 47-53, 2002.
**Jeffreys DA, and Axford JG**. Source locations of pattern-specific components of human visual evoked potentials. I. Component of striate cortical origin. *Exp Brain Res* 16: 1-21, 1972.
**Jewett DL, and Williston JS**. Auditory-evoked far fields averaged from the scalp of humans. *Brain* 94: 681-696, 1971.
**Kanno A, Nakasato N, Murayama N, and Yoshimoto T**. Middle and long latency peak sources in auditory evoked magnetic fields for tone bursts in humans. *Neurosci Lett* 293: 187-190, 2000.
**Keitel C, Andersen SK, and Muller MM**. Competitive effects on steady-state visual evoked potentials with frequencies in- and outside the alpha band. *Exp Brain Res* 205: 489-495, 2010.
**Kileny P, Paccioretti D, and Wilson AF**. Effects of cortical lesions on middle-latency auditory evoked responses (MLR). *Electroencephalogr Clin Neurophysiol* 66: 108-120, 1987.
**Korczak P, Smart J, Delgado R, Strobel TM, and Bradford C**. Auditory steady-state responses. *J Am Acad Audiol* 23: 146-170, 2012.
**Kraus N, McGee T, Carrell T, Sharma A, Micco A, and Nicol T**. Speech-evoked cortical potentials in children. *J Am Acad Audiol* 4: 238-248, 1993.
**Kraus N, Ozdamar O, Hier D, and Stein L**. Auditory middle latency responses (MLRs) in patients with cortical lesions. *Electroencephalogr Clin Neurophysiol* 54: 275-287, 1982.
**Krishnan A**. Human frequency following response. In: *Auditory evoked potentials: basic principles and clinical application*, edited by Burkard RF, Don JJ, and Eggermont JJ. Baltimore: Lippincott Williams & Wilkins, 2007, p. 315-335.
**Krishnan A, Bidelman GM, Smalt CJ, Ananthakrishnan S, and Gandour JT**. Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia* 50: 2849-2859, 2012.
**Krizman JL, Skoe E, and Kraus N**. Stimulus rate and subcortical auditory processing of speech. *Audiol Neurootol* 15: 332-342, 2010.
**Lee YS, Lueders H, Dinner DS, Lesser RP, Hahn J, and Klem G**. Recording of auditory evoked potentials in man using chronic subdural electrodes. *Brain* 107 ( Pt 1): 115-131, 1984.
**Liegeois-Chauvel C, Musolino A, Badier JM, Marquis P, and Chauvel P**. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol* 92: 204-214, 1994.
**Lopez-Calderon J, and Luck SJ**. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci* 8: 213, 2014.
**Luck SJ**. *An Introduction to the Event-Related Potential Technique*. Cambridge, Massachusetts; London, England: The MIT Press, 2014.
**Luck SJ, Kappenman ES. (Eds.)** *Oxford Library of Psychology. The Oxford Handbook of Event-Related Potential Components.* New York, NY: Oxford University Press, 2011.

**Mahajan Y, Davis C, and Kim J**. Attentional modulation of auditory steady-state responses. *PLoS One* 9: e110902, 2014.
**Makela JP, and Hari R**. Evidence for cortical origin of the 40 Hz auditory evoked response in man. *Electroencephalogr Clin Neurophysiol* 66: 539-546, 1987.
**Mangun GR, Hillyard SA, and Luck SJ**. Electrocortical substrates of visual selective attention. In: *Attention and performance XIV (silver jubilee volume)*, edited by David EM, and Sylvan KMIT Press, 1993, p. 219-243.
**Mendel MI, and Goldstein R**. Stability of the early components of the averaged electroencephalic response. *J Speech Hear Res* 12: 351-361, 1969.
**Miller L, Moore IVB, and Bishop C**. (Patent Pending). Frequency-multiplexed speech-sound stimuli for hierarchical neural characterization of speech processing. United States: The Regents of the University of California, 2017.
https://patents.google.com/patent/US20170196519A1/en
**Milner BA, Regan D, and Heron JR**. Differential diagnosis of multiple sclerosis by visual evoked potential recording. *Brain* 97: 755-772, 1974.
**Moller AR, Jannetta P, and Moller MB**. Intracranially recorded auditory nerve response in man. New interpretations of BSER. *Arch Otolaryngol* 108: 77-82, 1982.
**Moller AR, Jannetta PJ, and Moller MB**. Neural generators of brainstem evoked potentials. Results from human intracranial recordings. *Ann Otol Rhinol Laryngol* 90: 591-596, 1981.
**Moore JK**. The human auditory brain stem as a generator of auditory evoked potentials. *Hear Res* 29: 33-43, 1987.
**Morgan ST, Hansen JC, and Hillyard SA**. Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci U S A* 93: 4770-4774, 1996.
**Muller MM, Malinowski P, Gruber T, and Hillyard SA**. Sustained division of the attentional spotlight. *Nature* 424: 309-312, 2003.
**Muller N, Schlee W, Hartmann T, Lorenz I, and Weisz N**. Top-down modulation of the auditory steady-state response in a task-switch paradigm. *Front Hum Neurosci* 3: 1, 2009.
**Musacchia G, Strait D, and Kraus N**. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear Res* 241: 34-42, 2008.
**Naatanen R, and Picton T**. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24: 375-425, 1987.
**Norcia AM, Appelbaum LG, Ales JM, Cottereau BR, and Rossion B**. The steady-state visual evoked potential in vision research: A review. *J Vis* 15: 4, 2015.
**Pelizzone M, Hari R, Makela JP, Huttunen J, Ahlfors S, and Hamalainen M**. Cortical origin of middle-latency auditory evoked responses in man. *Neurosci Lett* 82: 303-307, 1987.
**Perrault N, and Picton TW**. Event-related potentials recorded from the scalp and nasopharynx. I. N1 and P2. *Electroencephalogr Clin Neurophysiol* 59: 177-194, 1984.
**Picton TW, Alain C, Woods DL, John MS, Scherg M, Valdes-Sosa P, Bosch-Bayard J, and Trujillo NJ**. Intracerebral sources of human auditory-evoked potentials. *Audiol Neurootol* 4: 64-79, 1999.
**Picton TW, Hillyard SA, Galambos R, and Schiff M**. Human auditory attention: a central or peripheral process? *Science* 173: 351-353, 1971.
**Picton TW, Hillyard SA, Krausz HI, and Galambos R**. Human auditory evoked potentials. I. Evaluation of components. *Electroencephalogr Clin Neurophysiol* 36: 179-190, 1974.
**Picton TW, John MS, Dimitrijevic A, and Purcell D**. Human auditory steady-state responses. *Int J Audiol* 42: 177-219, 2003.
**Picton TW, Woods DL, and Proulx GB**. Human auditory sustained potentials. I. The nature of the response. *Electroencephalogr Clin Neurophysiol* 45: 186-197, 1978a.

**Picton TW, Woods DL, and Proulx GB**. Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalogr Clin Neurophysiol* 45: 198-210, 1978b.
**Ponton C, Eggermont JJ, Khosla D, Kwong B, and Don M**. Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. *Clin Neurophysiol* 113: 407-420, 2002.
**Pratt H, and Sohmer H**. Intensity and rate functions of cochlear and brainstem evoked responses to click stimuli in man. *Arch Otorhinolaryngol* 212: 85-92, 1976.
**Regan D**. Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology* 20: 238-248, 1966.
**Regan D**. Steady-state evoked potentials. *J Opt Soc Am* 67: 1475-1489, 1977.
**Regan D, and Heron JR**. Clinical investigation of lesions of the visual pathway: a new objective technique. *J Neurol Neurosurg Psychiatry* 32: 479-483, 1969.
**Ross B, Picton TW, and Pantev C**. Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hear Res* 165: 68-84, 2002.
**Rupp A, Uppenkamp S, Gutschalk A, Beucker R, Patterson RD, Dau T, and Scherg M**. The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans(1). *Hear Res* 174: 19-31, 2002.
**Scherg M, and Von Cramon D**. Evoked dipole source potentials of the human auditory cortex. *Electroencephalogr Clin Neurophysiol* 65: 344-360, 1986.
**Scherg M, and Von Cramon D**. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalogr Clin Neurophysiol* 62: 32-44, 1985.
**Schoonhoven R, Boden CJ, Verbunt JP, and de Munck JC**. A whole head MEG study of the amplitude-modulation-following response: phase coherence, group delay and dipole source analysis. *Clin Neurophysiol* 114: 2096-2106, 2003.
**Sculthorpe-Petley L, Liu C, Hajra SG, Parvar H, Satel J, Trappenberg TP, Boshra R, and D'Arcy RC**. A rapid event-related potential (ERP) method for point-of-care evaluation of brain function: development of the Halifax Consciousness Scanner. *J Neurosci Methods* 245: 64-72, 2015.
**Shahin AJ, Roberts LE, Miller LM, McDonald KL, and Alain C**. Sensitivity of EEG and MEG to the N1 and P2 auditory evoked responses modulated by spectral complexity of sounds. *Brain Topogr* 20: 55-61, 2007.
**Shiga T, Althen H, Cornella M, Zarnowiec K, Yabe H, and Escera C**. Deviance-Related Responses along the Auditory Hierarchy: Combined FFR, MLR and MMN Evidence. *PLoS One* 10: e0136794, 2015.
**Shore SE, and Nuttall AL**. High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. *J Acoust Soc Am* 78: 1286-1295, 1985.
**Skoe E, and Kraus N**. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31: 302-324, 2010.
**Slugocki C, Bosnyak D, and Trainor LJ**. Simultaneously-evoked auditory potentials (SEAP): A new method for concurrent measurement of cortical and subcortical auditory-evoked activity. *Hear Res* 345: 30-42, 2017.
**Soustiel JF, Hafner H, Chistyakov AV, Yarnitzky D, Sharf B, Guilburd JN, and Feinsod M**. Brain-stem trigeminal and auditory evoked potentials in multiple sclerosis: physiological insights. *Electroencephalogr Clin Neurophysiol* 100: 152-157, 1996.
**Stapells DR, Linden D, Suffield JB, Hamel G, and Picton TW**. Human auditory steady state potentials. *Ear Hear* 5: 105-113, 1984.
**Starr A, and Hamilton AE**. Correlation between confirmed sites of neurological lesions and abnormalities of far-field auditory brainstem responses. *Electroencephalogr Clin Neurophysiol* 41: 595-608, 1976.
**Suzuki T, Hirai Y, and Horiuchi K**. Auditory brain stem responses to pure tone stimuli. *Scand Audiol* 6: 51-56, 1977.

**Tateuchi T, Itoh K, and Nakada T**. Neural mechanisms underlying the orienting response to subject's own name: an event-related potential study. *Psychophysiology* 49: 786-791, 2012.
**Thorpe SG, Nunez PL, and Srinivasan R**. Identification of wave-like spatial structure in the SSVEP: comparison of simultaneous EEG and MEG. *Stat Med* 26: 3911-3926, 2007.
**Tiitinen H, Sinkkonen J, Reinikainen K, Alho K, Lavikainen J, and Naatanen R**. Selective attention enhances the auditory 40-Hz transient response in humans. *Nature* 364: 59-60, 1993.
**Van Der Tweel LH, and Lunel HF**. Human Visual Responses to Sinusoidally Modulated Light. *Electroencephalogr Clin Neurophysiol* 18: 587-598, 1965.
**Vanni S, Warnking J, Dojat M, Delon-Martin C, Bullier J, and Segebarth C**. Sequence of pattern onset responses in the human visual areas: an fMRI constrained VEP source analysis. *Neuroimage* 21: 801-817, 2004.
**Vaughan HG, Jr., and Ritter W**. The sources of auditory evoked responses recorded from the human scalp. *Electroencephalogr Clin Neurophysiol* 28: 360-367, 1970.
**Vialatte FB, Maurice M, Dauwels J, and Cichocki A**. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Prog Neurobiol* 90: 418-438, 2010.
**Weber BA, and Folsom RC**. Brainstem wave V latencies to tone pip stimuli. *J Am Audiol Soc* 2: 182-184, 1977.
**Woldorff M, Hansen JC, and Hillyard SA**. Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. *Electroencephalogr Clin Neurophysiol Suppl* 40: 146-154, 1987.
**Woldorff MG, and Hillyard SA**. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol* 79: 170-191, 1991.
**Woods DL, Alain C, Covarrubias D, and Zaidel O**. Frequency-related differences in the speed of human auditory processing. *Hear Res* 66: 46-52, 1993.
**Yoshiura T, Ueno S, Iramina K, and Masuda K**. Human middle latency auditory evoked magnetic fields. *Brain Topogr* 8: 291-296, 1996.
**Yoshiura T, Ueno S, Iramina K, and Masuda K**. Source localization of middle latency auditory evoked magnetic fields. *Brain Res* 703: 139-144, 1995.
**Zhu L, Bharadwaj H, Xia J, and Shinn-Cunningham B**. A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. *J Acoust Soc Am* 134: 384-395, 2013.

## Tables

**Table 1. ERP Results.** Summary of single-subject results and group peak measurements (amplitude and latency) for each component. The "number of subjects" column indicates the number of subjects (out of 25) that showed a significant response for each component (using a p-threshold of 0.001). For the sustained components, the reported amplitude reflects an average across time (~225-425 ms for LLR Sustained Negativity; ~420-500 ms for VEP Late Negativity), instead of the peak amplitude.

| ERP Response | Component | Number of Subjects | Peak Amplitude: Mean (μV) | Peak Amplitude: Standard Error (μV) | Peak Latency: Mean (ms) | Peak Latency: Standard Error (ms) |
|---|---|---|---|---|---|---|
| ABR (Cz) | $n_0$ | 24 | -0.25 | 0.015 | 3.45 | 0.086 |
| | Wave V | 25 | +0.40 | 0.024 | 6.51 | 0.076 |
| | $N_0$ | 25 | -0.29 | 0.014 | 8.95 | 0.098 |
| | $P_0$ | 16 | +0.12 | 0.015 | 13.0 | 0.105 |
| MLR (Fz) | Na | 22 | -0.44 | 0.047 | 17.8 | 0.29 |
| | Pa | 17 | +0.23 | 0.046 | 25.0 | 0.31 |
| | Nb | 17 | -0.23 | 0.052 | 33.4 | 0.43 |
| | Pb | 20 | +0.31 | 0.042 | 44.1 | 0.32 |
| LLR (Fz) | P1 | 25 | +1.77 | 0.11 | 80 | 2.3 |
| | N1 | 21 | -0.66 | 0.11 | 173 | 2.6 |
| | Sustained Negativity | 23 | -0.64 (average) | 0.08 (SE of average) | 323 | 11.4 |
| | | | | | | |
| VEP (Oz) | P1 | 22 | +3.65 | 0.39 | 106 | 2.3 |
| | N1 | 17 | -2.31 | 0.47 | 177 | 3.0 |
| | P2 | 20 | +2.54 | 0.45 | 250 | 2.7 |
| | Late Negativity | 20 | -1.62 (average) | 0.31 (SE of average) | 468 | 4.7 |

**Table 2. Steady-State Results.** Summary of single-subject results, as well as the group magnitudes and signal-to-noise ratios (SNRs) at each frequency of interest of the ASSR and SSVEP. The "number of subjects" column indicates the number of subjects (out of 25) that showed a significant response for each component (using Bonferroni-corrected p-thresholds). These results were derived from the magnitude spectra and noise floor estimates that were averaged in frequency space across epochs and draws, within the bootstrapping algorithm.

| Steady-State Response | Frequency (Hz) | Number of Subjects | Raw Magnitude: Mean (dB, arbitrary units) | Raw Magnitude: Standard Error (dB, arbitrary units) | SNR: Mean (dB) | SNR: Standard Error (dB) |
|---|---|---|---|---|---|---|
| ASSR (Fz) | 41 | 25 | -18.8 | 0.90 | 16.4 | 0.82 |
| | 82 | 24 | -25.6 | 1.18 | 12.7 | 1.22 |
| | 123 | 25 | -27.6 | 0.67 | 13.2 | 0.63 |
| | 164 | 25 | -28.5 | 0.55 | 14.7 | 0.62 |
| SSVEP (Oz) | 7.5 | 21 | -11.5 | 0.65 | 3.8 | 0.50 |
| | 12 | 24 | -13.5 | 0.64 | 3.4 | 0.43 |
| | 15 | 22 | -10.9 | 1.08 | 7.3 | 0.90 |
| | 24 | 19 | -17.4 | 0.91 | 4.1 | 0.67 |

## Figure Captions

**Figure 1. Stimuli and Experimental Design. A**) Examples of the Black/White (left) and Red/Green (right) stimuli are displayed. **B**) A spectrogram of the Cheech is shown on the left for one of the sentences. On the right, a zoomed-in view of a chirp train is illustrated, to demonstrate that the second chirp in each train was omitted for recording a clear auditory MLR. **C**) An overview of the stimulus presentation is depicted.

**Figure 2. ABR Results. A**) Single-subject ABRs in channel Cz are shown, thresholded at $p < 0.001$, with non-significant data samples set to an amplitude of 0 μV. **B**) Group-average ABRs derived from the bootstrapping procedure are shown. The top panel shows the group-average time waveform, while the group-average scalp topographies of the significant deflections are displayed below. In this figure and in Figures 3, 4, and 5, the gray shaded box in the group time waveform plot depicts the group average (plus standard error of the mean) range of amplitudes that were not significant, using a p-threshold of 0.001.

**Figure 3. MLR Results. A**) Single-subject MLRs in channel Fz are shown, thresholded at $p < 0.001$, with non-significant data samples set to an amplitude of 0 μV. **B**) The group-average MLR time waveform is shown, along with the scalp topographies of the significant MLR components.

**Figure 4. LLR Results. A**) Single-subject LLRs in channel Fz are shown, thresholded at $p < 0.001$, with non-significant data samples set to an amplitude of 0 μV. **B**) The group-average LLR time waveform is displayed, along with the scalp topographies of the auditory P1 and N1/sustained negativity.

**Figure 5. VEP Results. A**) Single-subject VEPs in channel Oz are displayed, which have been thresholded at $p < 0.001$, with non-significant data samples set to an amplitude of 0 μV. **B**) The

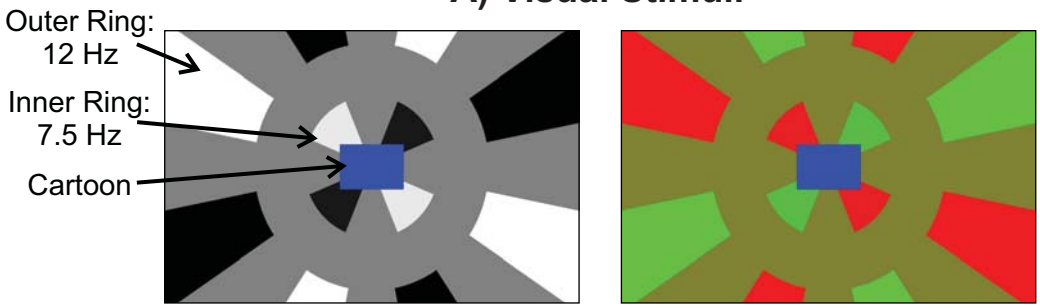
group-average VEP time waveform is shown, along with the scalp topographies of the significant VEP components

**Figure 6. ASSR Results. A**) Single-subject data, indicating signal-to-noise ratio of the 41 Hz ASSR and the first three harmonics (82, 123, and 164 Hz), are displayed, for channel Fz. Non-significant responses were set to 0 dB; note that only one participant had a non-significant ASSR: Subject 16 at 82 Hz. **B**) Group-average raw ASSR magnitude and the signal-to-noise ratio, along with the noise floor estimate, are depicted; these responses were created by averaging data in the frequency domain. Additionally, the raw magnitude computed by averaging data in the time domain before converting to the frequency domain, is plotted. Below are the group-average scalp topographies of the raw ASSR magnitude (averaged in the frequency domain) at the four frequencies of interest. Abbreviations: SNR, signal-to-noise ratio; n.s., not significant.

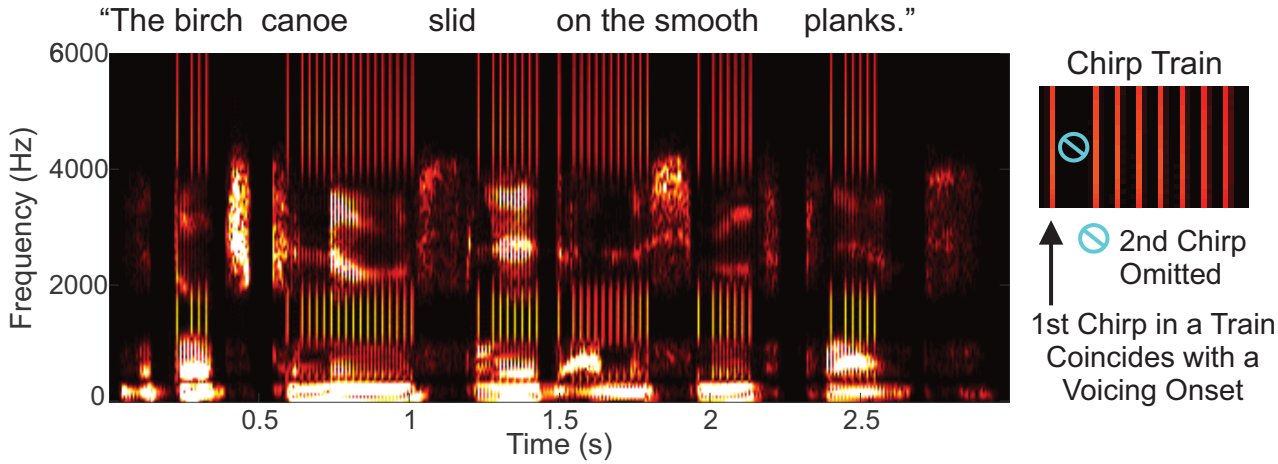
**Figure 7. SSVEP Results. A**) Single-subject responses, indicating signal-to-noise ratio at 7.5 Hz (inner ring), 12 Hz (outer ring), 15 Hz (inner ring harmonic), and 24 Hz (outer ring harmonic) are displayed, for channel Oz. Non-significant SSVEP responses were set to 0 dB. **B**) Group-average raw SSVEP magnitude and the signal-to-noise ratio, along with the noise floor estimate, are depicted; these responses were created by averaging data in the frequency domain. Additionally, the raw magnitude computed by averaging data in the time domain before converting to the frequency domain, is plotted. Below are the group-average scalp topographies of the raw SSVEP magnitude (averaged in the frequency domain) at the four frequencies of interest. Abbreviations: SNR, signal-to-noise ratio; n.s., not significant.

**Figure 8. Z-Scored Single-Subject Data.** Aggregate z-scores for each EEG response are plotted for each subject.

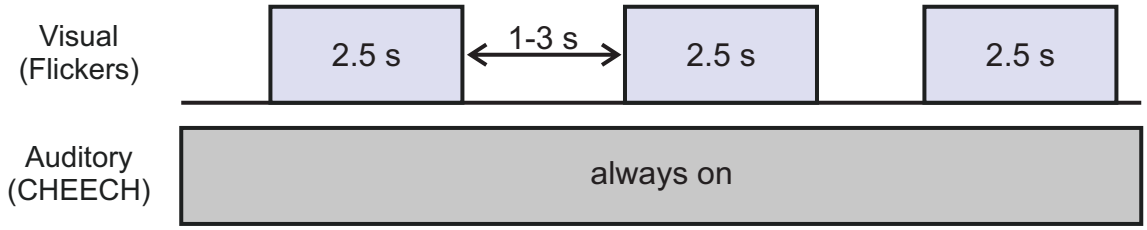
## A) Visual Stimuli

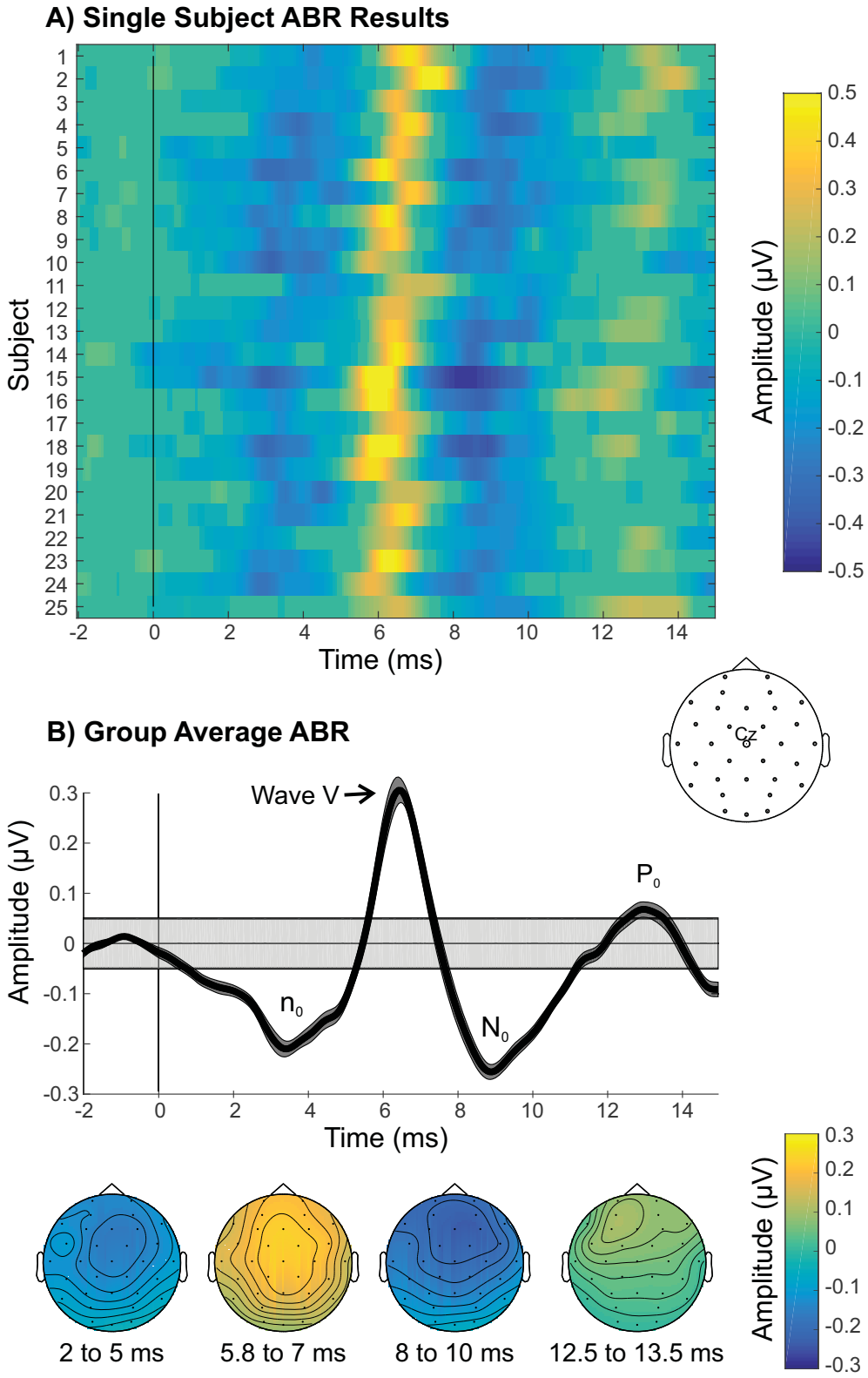


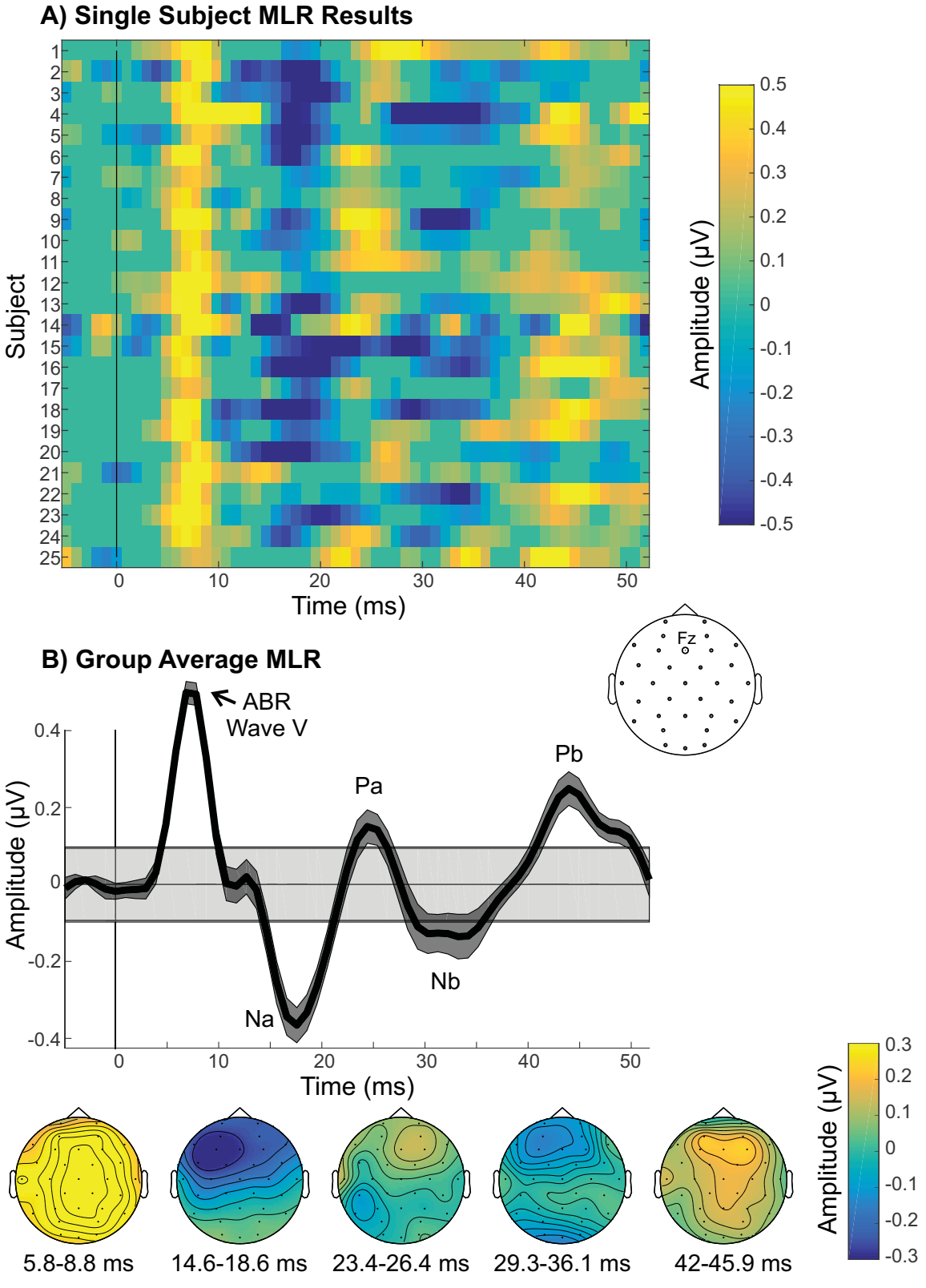

## B) Auditory Stimuli: CHEECH (CHirp-spEECH)

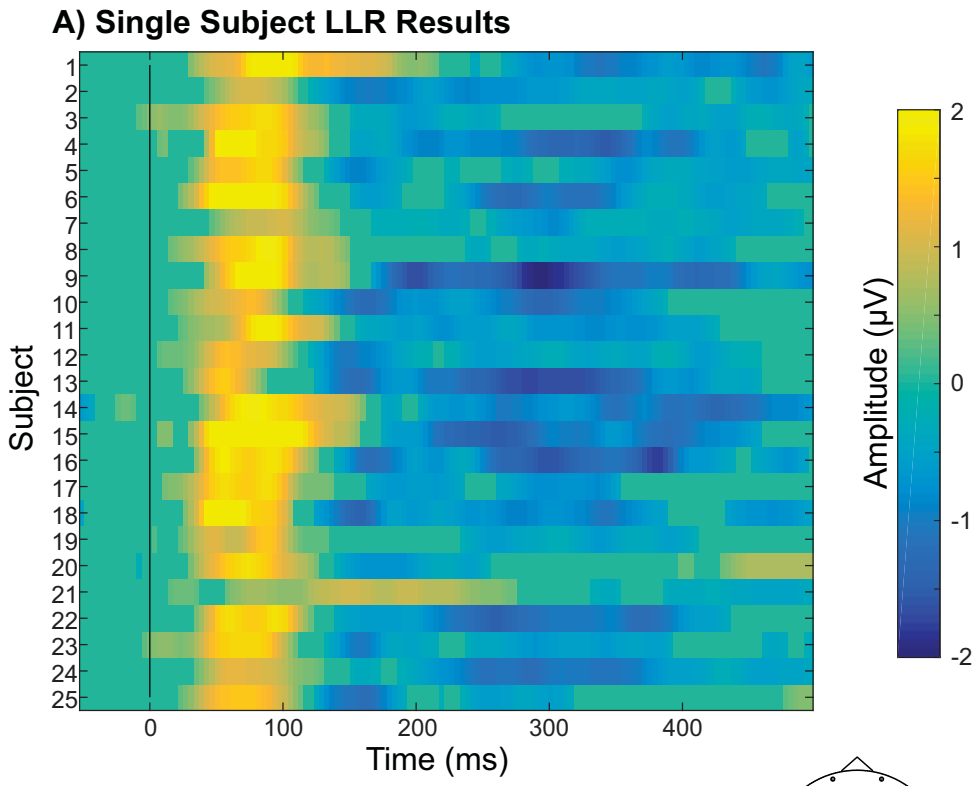


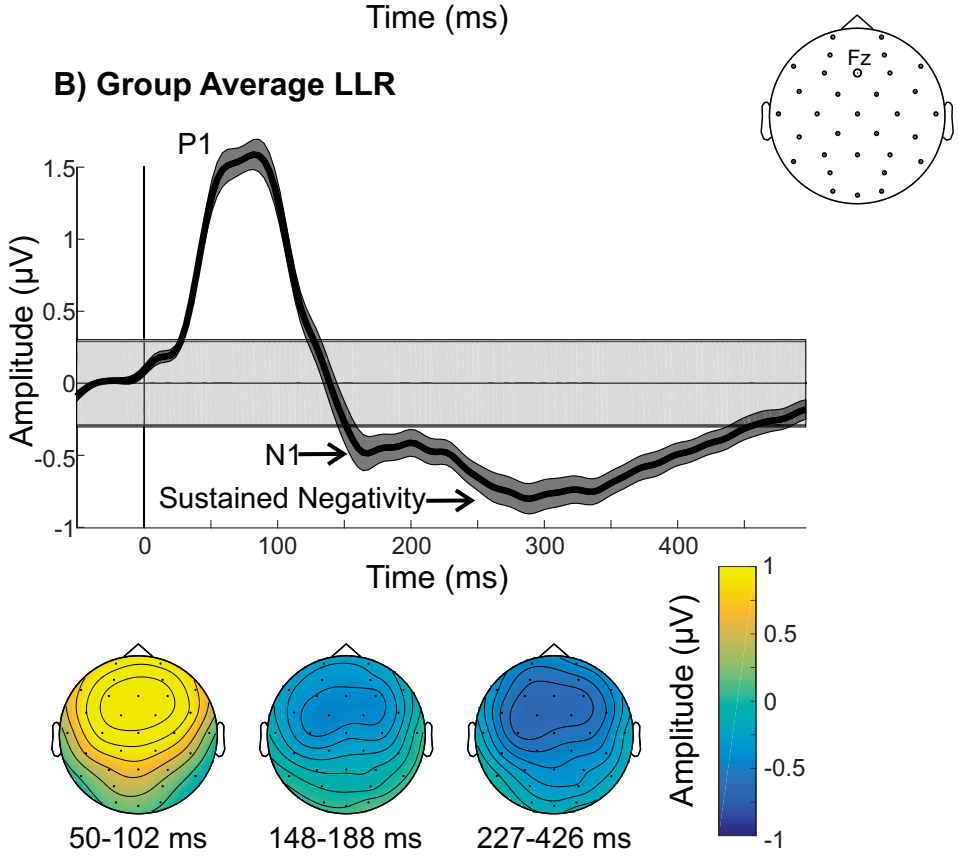

## C) Overview of Stimulus Presentation

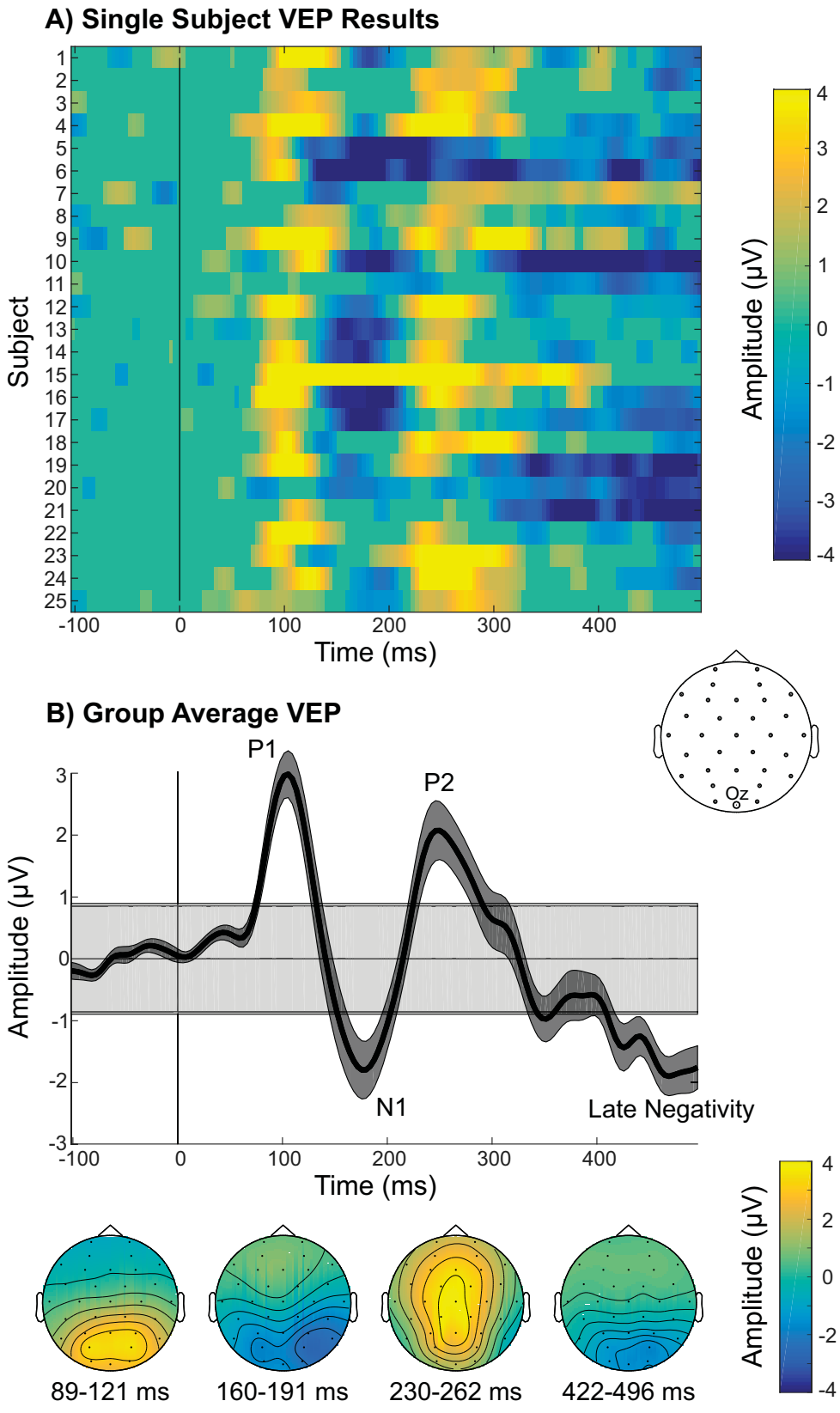


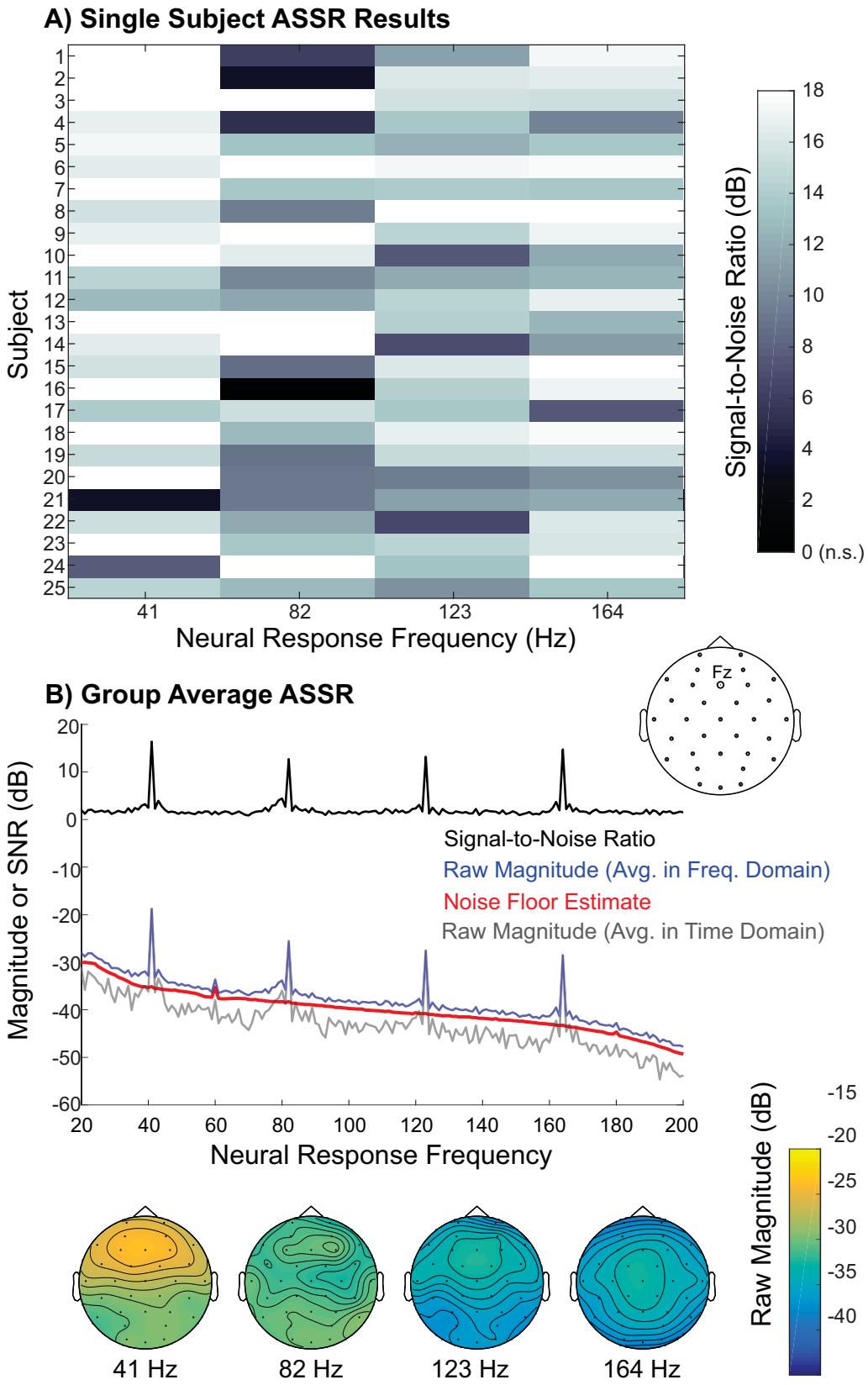

A) Single Subject ABR Results
Subject
Time (ms)
Amplitude (μV)
B) Group Average ABR
Wave V
$P_0$
$n_0$
$N_0$
Amplitude (μV)
Time (ms)
Cz
2 to 5 ms
5.8 to 7 ms
8 to 10 ms
12.5 to 13.5 ms
Amplitude (μV)

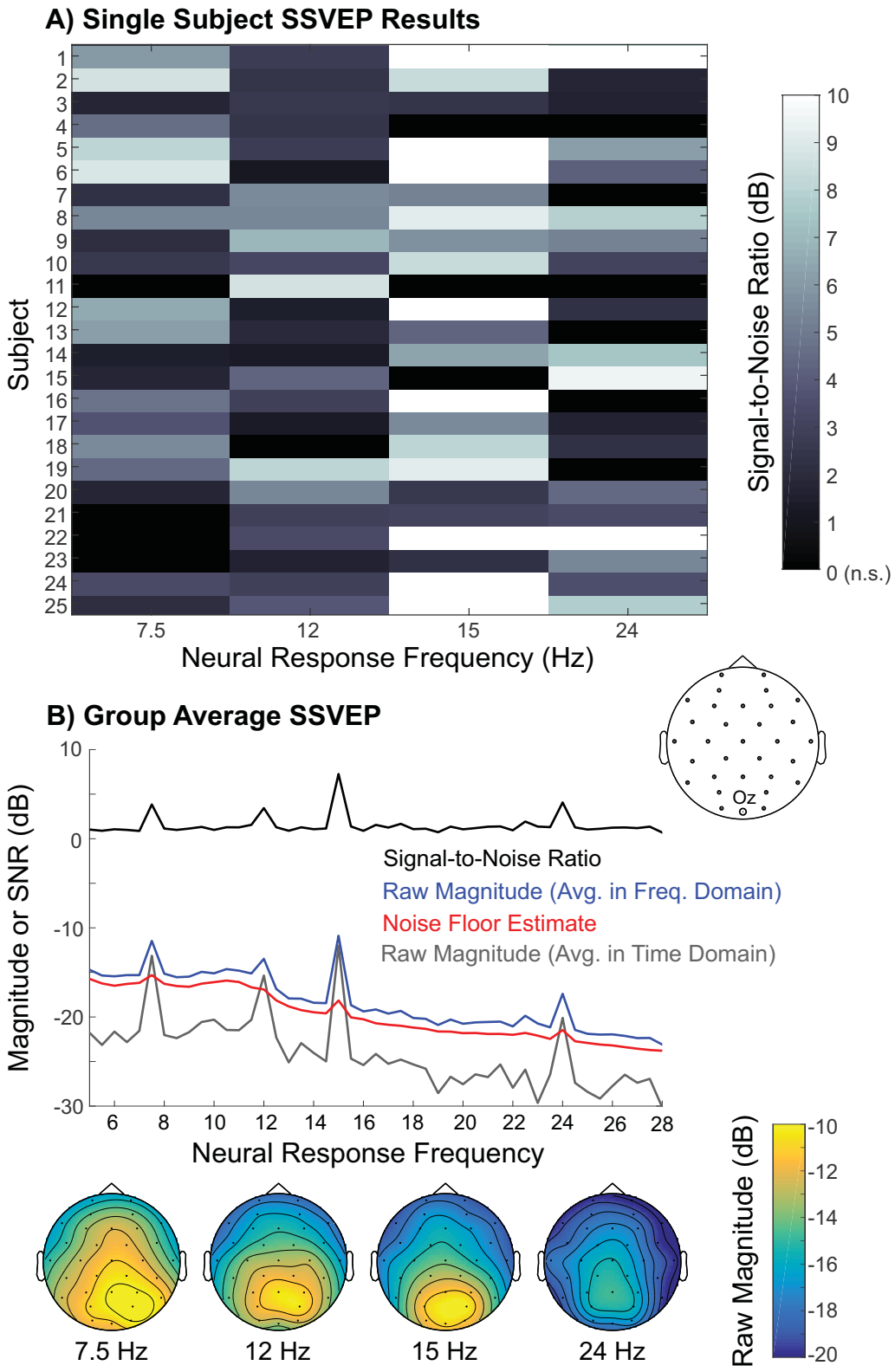

A) Single Subject MLR Results
Subject
Time (ms)
Amplitude (μV)
B) Group Average MLR
ABR
Wave V
Pa
Pb
Na
Nb
Fz
Amplitude (μV)
Time (ms)
5.8-8.8 ms
14.6-18.6 ms
23.4-26.4 ms
29.3-36.1 ms
42-45.9 ms
Amplitude (μV)

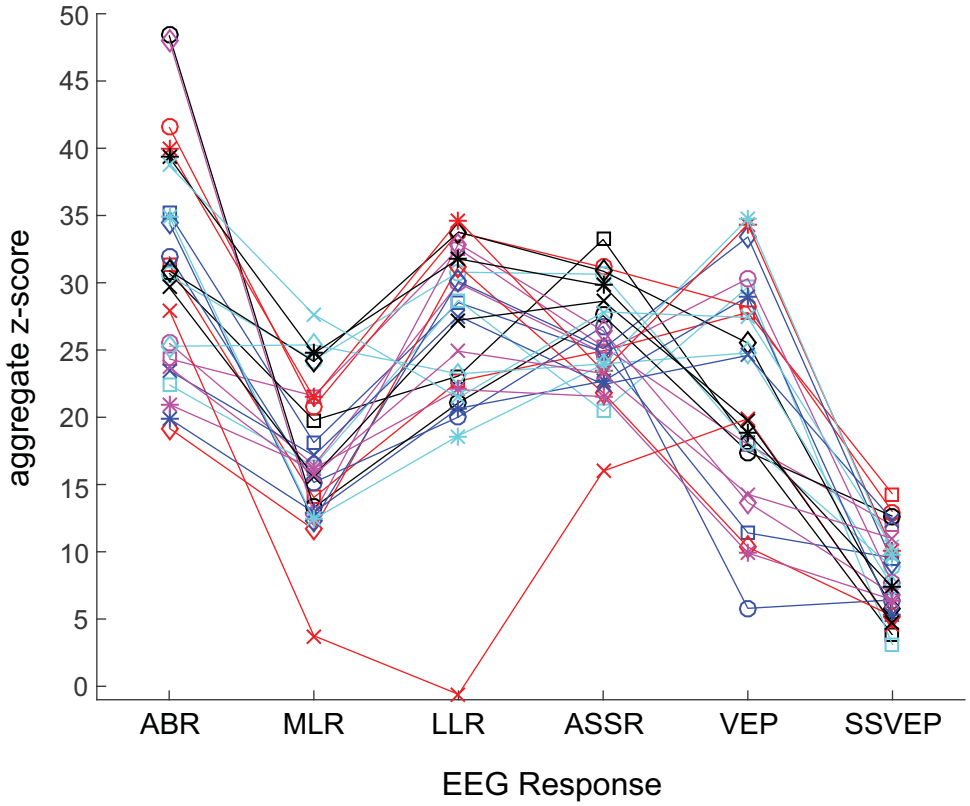

A) Single Subject LLR Results
Subject
Time (ms)
Amplitude (μV)
2
1
0
-1
-2


B) Group Average LLR
P1
N1
Sustained Negativity
Amplitude (μV)
Time (ms)
Fz
50-102 ms
148-188 ms
227-426 ms
Amplitude (μV)


A) Single Subject VEP Results
Subject
Time (ms)
Amplitude (μV)
B) Group Average VEP
P1
P2
N1
Late Negativity
Oz
Amplitude (μV)
Time (ms)
89-121 ms
160-191 ms
230-262 ms
422-496 ms
Amplitude (μV)


A) Single Subject ASSR Results
Subject
Neural Response Frequency (Hz)
Signal-to-Noise Ratio (dB)
B) Group Average ASSR
Fz
Magnitude or SNR (dB)
Signal-to-Noise Ratio
Raw Magnitude (Avg. in Freq. Domain)
Noise Floor Estimate
Raw Magnitude (Avg. in Time Domain)
Neural Response Frequency
Raw Magnitude (dB)
41 Hz
82 Hz
123 Hz
164 Hz


A) Single Subject SSVEP Results
Subject
Neural Response Frequency (Hz)
7.5
12
15
24
Signal-to-Noise Ratio (dB)
10
9
8
7
6
5
4
3
2
1
0 (n.s.)
B) Group Average SSVEP
Oz
Magnitude or SNR (dB)
Signal-to-Noise Ratio
Raw Magnitude (Avg. in Freq. Domain)
Noise Floor Estimate
Raw Magnitude (Avg. in Time Domain)
Neural Response Frequency
Raw Magnitude (dB)
7.5 Hz
12 Hz
15 Hz
24 Hz


aggregate z-score
50
45
40
35
30
25
20
15
10
5
0
ABR
MLR
LLR
ASSR
VEP
SSVEP
EEG Response