

Differences in Click and Speech Auditory Brainstem Responses and Cortical Response Patterns: A Pilot Study

Jean-François Knebel^{1,2}, Arnaud Jeanvoine¹, Fabian Guignard^{1,3}, Jean-Marc Vesin⁴ and Céline Richard^{1,3*}

¹Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology and Department of Clinical Neurosciences, University Hospital Center and University of Lausanne, Lausanne, Switzerland

²EEG Brain Mapping Core Centre for Biomedical Imaging (CIBM), Lausanne, Switzerland

³Head and Neck Surgery Department, Lausanne University Hospital (CHUV), Lausanne, Switzerland

⁴Applied Signal Processing Group, Swiss Federal Institute of Technology, Lausanne, Switzerland

Abstract

Hearing loss has adverse developmental, cognitive and social impacts on individual's lives. To improve its diagnosis and remediation outcomes, development of objective hearing methods offer new ways of optimizing care strategy. Complementary to the classical click-evoked brainstem and cortical responses, interest grows regarding the speech auditory brainstem responses, whose components, namely the onset and frequency following responses, are proposed as biomarkers of speech encoding at the brainstem level. To improve the understanding of auditory processing in the human brainstem and its effect on cortical processing, we used a multimodal set-up and recorded brainstem and cortical potentials in response to click *stimuli* and speech *stimuli* presented at different stimulation intensities in normal hearing adults. We hypothesized that even though click- and onset responses of the speech-auditory brainstem responses share some similarities; their underlying mechanisms are in some point distinct. We also presumed that using a noninvasive method we could assess the generators underlying click, onset and frequency following responses and the time wise influence of brainstem encoding on cortical processing of click and speech *stimuli*. Results showed evidence regarding mechanisms underlying onset response and frequency following response components at the brainstem level. A direct time wise relationship between subcortical encoding and cortical encoding was revealed: cortical activity in the left cortex was related to the onset response latency at the brainstem level. These results highlight the potential use of new methods in speech processing electroencephalographic studies and provide qualitative and topographical data regarding subcortical and cortical auditory processing network.

Keywords: Click-evoked auditory brainstem responses; Speech-evoked auditory brainstem responses; Cortical auditory responses; Subcortical generators; Cortical generators

Introduction

How speech and non-speech sounds are processed at the subcortical and cortical levels is relatively poorly understood. This knowledge is important for evaluating hearing impairment and providing patients with optimal rehabilitation strategies and communication training, especially for young children or people unable to provide reliable feedback on their hearing experience. A first step toward achieving this goal is to elucidate the biological mechanisms that underlie auditory processing. The currently available methods for studying these mechanisms include both subjective and objective techniques with varying degrees of invasiveness. Currently, auditory evoked responses represent the optimal compromise for a thorough assessment that is also noninvasive. Auditory evoked responses stem from the neural activity generated by subcortical regions located in the brainstem (auditory brainstem responses, ABRs), the thalamus (middle latency responses), and also cortical generators (cortical auditory evoked potentials, CAEPs) [1,2]. Analyses of the response morphology of ABRs and CAEPs offer a temporal window to noninvasively observe the neural representation of speech processing and how the subcortical and central auditory mechanisms interact.

For years, auditory evoked potentials have been used in children and adults to explore in a noninvasive, reliable manner the neural transmission of various types of *stimuli*, including clicks [3], chirps and tone-bursts [4-6]; steady-state signals such as amplitude modulated (AM) tones in healthy hearing [7,8] and rehabilitated patients [9] have also been employed. While these *stimuli* can be easily implemented in a clinical setting, they do not reflect the complex nature of the information transmitted and integrated by the auditory system during

traditional daily communication [10,11]. Toward bridging this gap in both research and clinical communities, using speech-elicited auditory brainstem responses (speech ABRs) has become of growing interest [12-16]. Speech ABR is an objective, noninvasive electrophysiological approach for studying auditory neural coding at the brainstem level [13,14,17]. This neural response to speech includes both a transient response (onset response, OR) to the non-periodic part of the stimulus and a sustained phase-locked response (frequency following response, FFR) to the periodic portions [18,19]. Importantly, the FFR has been found to be highly replicable [20,21] and can provide robust biomarkers of auditory processing at the brainstem level in humans [13,14,19,22-25] as well as top-down interactivity of the auditory system [17,26]. These reports emphasize the role of the brainstem as a hub of interconnected ascending and descending pathways, prone to neural adaptation in response to learning [27]. Studies have focused on the relationship between subcortical encoding and CAEPs [28] and their potential integration into clinical practice [29,30]. Despite great strides toward understanding the anatomical and functional organization of the auditory brainstem system, its interconnectivity with cortical structures remains only partially understood. The mechanisms cited

*Corresponding author: Céline Richard, ENT, Head and Neck Surgery Department, Lausanne University Hospital (CHUV), Lausanne, Switzerland, Tel: 41 79 556 56 09; E-mail: celine.richard@chuv.ch

Received April 06, 2018; Accepted May 07, 2018; Published May 12, 2018

Citation: Knebel JF, Jeanvoine A, Guignard F, Vesin JM, Richard C (2018) Differences in Click and Speech Auditory Brainstem Responses and Cortical Response Patterns: A Pilot Study. J Neurol Neurophysiol 9: 463. doi:10.4172/2155-9562.1000463

Copyright: © 2018 Knebel JF, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

for accurately encoding the many speech-related acoustic cues are still largely speculative. Indeed, controversies remain regarding the relationship between OR and click-elicited wave V, some arguing similar mechanisms at the subcortical level while others suggest different encoding processes [10,22]. In line with the uncertainty related to wave V of the click-evoked potentials and OR of the speech ABR generating mechanisms, the location of FFR generators in the subcortical levels is also still debated. Some reports suggest the FFR emerges from regions lower than the inferior colliculus [31,32], while others propose a strong contribution from the inferior colliculus [33-36]) supported by FFR latency analyses [37,38]. Previously, a magneto-encephalographic study proposed an additional right hemispheric predominant contribution from the auditory cortex at the 100 Hz FFR fundamental frequency [39]. Conversely, other recent reports favor the idea of the FFR representing a composite of activity from different sources in the auditory system [40]. Although one may argue that previous studies correlated speech ABR and cortical encoding [41], there is still a lack of knowledge regarding potential clinical use of EEG techniques to assess the topographical and qualitative relationship between click and speech ABR as well as the connectivity between brainstem and cortical speech auditory potentials.

We have developed methodology based on the high temporal accuracy of a multichannel EEG system and the information this generates allows for advanced processing and analysis methods that may be used in children. Therefore, we aimed to compare click- and speech processing as a function of intensity at both the brainstem and cortical level. We performed a direct investigation of the relationship between subcortical and cortical activity for the wave V of the click ABR, the OR, and the FFR of the speech ABR. Since we found interesting differences between the characteristics of the OR and FFR of the speech ABR and CAEP, we conducted an exploratory analysis in adult listeners using validated source modeling techniques [42] to identify their underlying generators in the brainstem.

Materials and Methods

Participants

Eight French native young adult speakers (mean age: 24.7 years, SD: 0.88 years) with similar educational levels participated in the present study. All participants provided signed informed consent documents prior to their enrollment. This study and its related methods were approved by the University Hospital Research Ethics Committee of Lausanne (#PB_2016-02008) and were performed in accordance with the ethical standards and good practices guidelines as put forth by the Declaration of Helsinki.

No participant had been diagnosed with a hearing, language, or neurological disorder. In order to avoid the influence of musical training on speech processing, participants with formal musical training were excluded [15]. All participants were strongly right-handed (>72% laterality) according to the Edinburgh Handedness Inventory [43]. Prior to inclusion, participants underwent clinical examination consisting of an otoscopy, otoacoustic emission test, and audiometric testing to ensure typical hearing thresholds from 125 to 8000 Hz on both sides and to determine pure-tone averages (PTAs) for each participant at 500, 1000, 2000 and 4000 Hz (mean=8.8; SD=3.3; max=15 dB HL).

Stimulation setting and *stimuli* characteristics

Setting: *Stimuli* were sent using a SoundBlaster Audigy^X-FI 5.1 Surround Sound Card and delivered to the insert earphones. To avoid any time delay between the recorded brainstem and cortical signals,

the soundcard was connected to a trigger that delivered a Transistor-transistor logic impulsion to the EEG recording system. For all EEG recordings, participants sat comfortably in an electromagnetically-shielded soundproofed room while watching a subtitled movie without a soundtrack [44]. In order to avoid any attention-induced modification of neural activation in the auditory cortex, participants were instructed not to focus on the sound [45].

We recorded ABRs to clicks and speech stimulus tokens at five intensities to better match settings commonly used in clinical practice and cortical responses at three intensities. To avoid any stimulus artifacts, ER-3A Insert Earphones (Etymotic Research, Elk Grove Village, IL, USA) were used. Auditory brainstem and cortical potentials were scalp-recorded separately in response to both click and speech *stimuli*. To emulate realistic conditions and obtain larger and more robust responses, binaural stimulation was used [41,44,46].

Click stimulation: Similar to conditions routinely used in clinics, clicks of 200 μ s with a repetition rate of 20/s were presented in alternate polarity [47]. These clicks were delivered binaurally through the insert earphones along a seven-step intensity continuum from 60 dB SPL to 0 dB SPL at the subcortical level and from 60 dB SPL to 30 dB SPL for cortical responses (according to the hearing threshold (dB HL) previously identified for each participant). A total of 2000 epochs were presented to the listeners for subcortical responses. Cortical responses were elicited from 300 epochs (alternate polarity) of the same click with a random jitter of 200-300 ms (to avoid α -band entrainment) and an average inter-stimulus interval (ISI) of 750 ms.

Speech stimulus: Given evidence illustrating the importance of using natural sound [48], we used a 202 ms length natural consonant-vowel (CV) /ba/ syllable (/b/=110 ms; F0: 200 Hz; F1: 750 Hz; F2: 1500 Hz) for both subcortical and cortical recordings. The CV syllable /ba/ was chosen based on both clinical evidence of adult phoneme perception's dependence on the subject's native language [49,50] and its common use in the French literature regarding speech ABR [38,51]. Moreover, in line with a possible future application in infants, the choice of a natural voice quite similar to a mother's seemed highly relevant. The /ba/ syllable was binaurally presented through ER 3A Insert Earphones (Etymotic Research, Elk Grove Village, IL, USA), 2000 epochs (alternate polarity to enable canceling of the cochlear microphonic [44]), 3.1/s with an ISI of 75 ms for subcortical stimulation and average ISI of 750 ms for cortical stimulation (jitter 200 ms). Stimulation intensities for both subcortical and cortical responses ranged from 30 to 60 dB SPL in 10 dB steps. The test presentation order was counter balanced across intensities and *stimuli*.

EEG Recording and Preprocessing

Recording: EEG data were recorded from 32 channels using an actiCHamp EEG recording system with actiCAP active electrodes (Brain Products GmbH, Germany) with electrode impedance kept below 25 k Ω (thereby preventing overly noisy recordings). EEG signals were referenced against Fz, amplified by an actiCHamp amplifier (Brain Products GmbH, Germany), sampled at 10 KHz and stored for offline analysis. In order to optimize recording length, brainstem and cortical-evoked responses were collected separately [44].

Pre-processing:

Subcortical potentials: Click- and speech-evoked potentials were obtained by averaging EEG epochs from -25 to 25 ms (click) and -25 to 300 ms (/ba/) post-stimulus onset. Traces were filtered between 80 and 2000 Hz (Butterworth filter, notch filter 50 Hz). Epochs with amplitude

deviations greater than $\pm 80 \mu\text{V}$ in any channel were considered artifacts and thus rejected. The traces were analyzed using an average reference [28] and a classical mastoid reference (mean mastoids [44]). Each run and recording included the responses to 2000 clicks (alternate polarity; 40 dB SPL: 3801 epochs (3148-3976), 50 dB SPL: 3800 (3465-3959), 60 dB SPL: 3894 (3696-3995]) or 2000 /ba/ (40 dB SPL: 3706 (3438-3881)), 50 dB SPL: 3636 (3292-3911), 60 dB SPL: 3719 (3495-3928)). Validity was statistically assessed: 6×1 one-way ANOVA $F(5, 42)=1.81$; $p=0.13$.

Brain potentials acquisition and pre-processing: Event-related potentials were obtained from 32 active electrodes (impedances $<25 \text{ k}\Omega$, Fz reference, 0.1-40 Hz bandpass filter, notch filter 50 Hz, 1000 Hz sampling rate). For auditory ERP calculation, EEG epochs were time-locked to the presentation of the sound and spanned 100 ms pre-stimulus and 500 ms post-stimulus. Epochs with amplitude deviation greater than $\pm 80 \mu\text{V}$ at any channel were considered artifacts and were rejected. Data from 'bad' channels were interpolated using 3D splines [52]. Prior to grand-averaging, data were re-calculated to an average reference and a baseline correction was applied using the 100 ms pre-stimulus period. For each participant, eight auditory ERPs were calculated following the two test conditions (/ba/ and clicks). The number of accepted sweeps per condition was (mean, range) /ba/ 40 dB SPL: 472 epochs (421-556), /ba/ 50 dB SPL: 463 (339-531), /ba/ 60 dB SPL: 483 (299-567)) and click 40 dB SPL: 478 (412-550)), click 50 dB SPL: 471 (354-530), click 60 dB SPL: 472 (286-570)). Validity was statistically assessed: 6×1 one-way ANOVA $F(5,42)=0.08$; $p=0.99$.

Data analysis

Brainstem evoked potentials: Experienced observers identified waves I-III and V for the click-evoked responses and waves V-A (peaks of the OR complex) for the speech ABRs of each subject and intensity. Observers were blinded to the test conditions. To ensure correct peak identification, OR and FFR latency were also measured using dynamic time warping [53] because the standard cross-correlation technique usually used at 60 dB SPL didn't provide reliable results at 40 dB SPL. Wave V of the click ABR, wave V latency, wave A latency, VA interpeak slopes, duration and amplitude (voltage difference) of these peaks for the speech ABR were measured. Source generators as well as intensity effects were evaluated as described in the following section.

Cortical analysis:

General overview: ERP analyses were performed using Cartool freeware (<http://sites.google.com/site/fbmlab/cartool/cartooldownload>), Python-based LINEViewer and STEN utilities (<http://unil.ch/line/home/menuinst/about-the-line/software-analysis-tools.html>). Effects were identified with an analysis procedure referred to as electrical neuroimaging [54,55], which allows for direct assessment of reference-independent global measures of the electric field at the scalp as well as distributed source estimations. Using these reference-independent global measure analyses, we were able to differentiate effects due to modulations in the strength of responses of statistically indistinguishable brain generators from alterations in the configuration of the active generators (inferred from the topography of the electric field at the scalp).

Global field power measures: Brain microstate variations have been proposed to reflect rapid switching between neural networks [56]. These variations are reflected in the brain's electric field configuration [57] and can be ascertained by calculating the global field power (GFP [58,59]). GFP acts as a reference-independent descriptor of the potential field, allowing for determination of component latency and topographical changes across all EEG electrodes as a function of

time [60]. The ERP components were extracted using peaks of GFP. Statistical analyses were performed on GFP values at each time point. The statistical design was a repeated measures 2×3 ANOVA using within-subject factor sounds (click, /ba/) and intensities (40, 50 and 60 dB SPL). Temporal auto-correlation at GFP levels was corrected through applying a 150 contiguous data-point temporal criterion (15 ms at 10 KHz sampling) for the persistence of differential effects [61].

Topography consistency testing: The topographical change occurring in subsequent potential field distributions was analyzed using global dissimilarity (DISS) [59]. The DISS is directly correlated to the spatial correlation coefficient and provides a measure of topographic instability between two electric fields. DISS values at each time point were compared with RAGU software [62] and non parametric repeated measures 2×3 ANOVA using within-subject factor sounds (click, /ba/) and intensities (40, 50 and 60 dB SPL). In addition to this time-based measure, a topographic consistency test (TCT) [63] was conducted across six conditions for each time point. Based on this analysis, we ascertained the consistency of the observed effect across subject and for each condition. To account for temporal auto-correlation, only effects ($p < 0.05$) persisting for at least 150 time frames ($>15 \text{ ms}$ at 10 KHz sampling) were considered reliable [61].

Source estimation: To estimate generator sources involved in click and /ba/ processing, we conducted source estimation analyses at both subcortical and cortical (ERP) levels. The LAURA algorithm was used to estimate the neural sources of the electric signal recorded at the 32-head surface active sensors (31 recording channels and one reference electrode) by using an inverse solution matrix consisting of 5104 nodes equally distributed within the grey matter of the Montreal Neurological Institute (MNI) average brain and generated with the Spherical Model with Anatomical Constraints (SMAC; [64]). For each subject, 4000 epochs were randomly chosen and processed to illustrate baseline activity. The activity of each node is provided in $\mu\text{A}/\text{mm}^2$ with a spatial accuracy of $6 \times 6 \times 6 \text{ mm}$ [65,66]. Only activities above the 95 percentile were used for source estimation identification. The time periods used for source estimations were determined for each subject as follows: the beginning and the end of the OR component, the beginning and the end of the FFR and to ensure the entire FFR was fit, a Fast Fourier Transform to localize the FFR F0. Then, source estimations were performed using the time period -1 ms and +1 ms including the wave V, as well as the P50 and P300 periods extracted from the GFP grand mean levels.

Results

Phase-locking activity and reproducibility of speech ABR versus click-evoked ABR

Click-elicited ABR revealed a well-defined and reproducible wave V until 0 dB SPL for all subjects. The OR and FFR components of the speech ABR were identified in the grand average of the individual neural responses and were found to be very well-defined when recorded at stimulation intensities from 60 dB SPL down to 40 dB SPL (Figure 1A). However, the FFR peak synchronization to the periodic part of the /ba/ was less reproducible at 30 dB SPL (peak's amplitude not above the pre-stimulus amplitude) and unrecognizable at lower intensities. This observation is consistent with previous reports documenting the brainstem response elicited by speech [38]. In contrast, speech ABR components were present and reproducible across participants from 60 to 40 dB SPL, with spectral features of the grand average phase locked up to the second harmonic, in accordance with brainstem phase-locking activity in brainstem nuclei (lower brainstem), inferior colliculus and

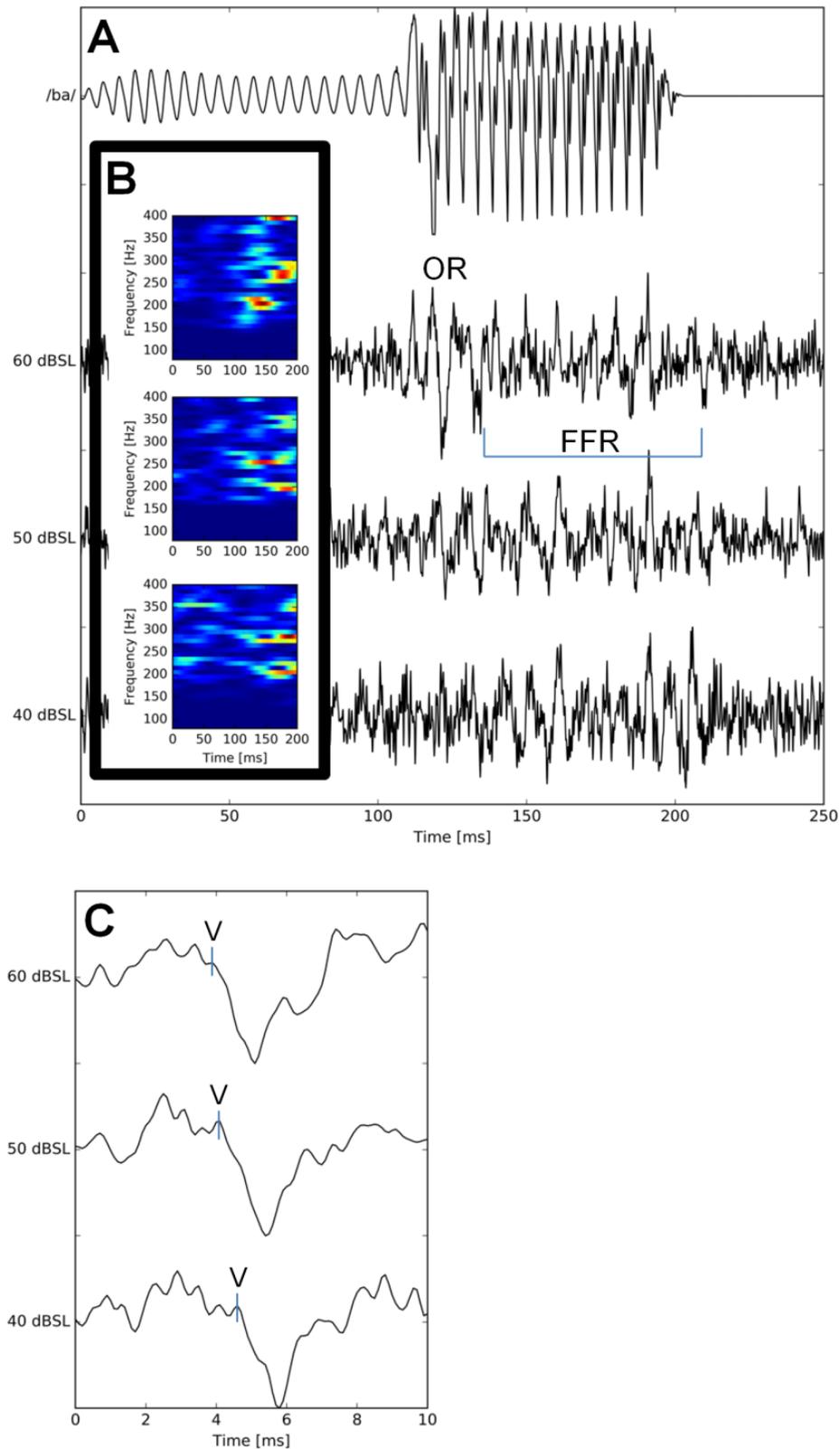


Figure 1: Single subject onset response (OR) and Frequency Following Response (FFR) components, in response to the /ba/ stimulus (upper row), were clearly identified on temporal representation from 40 to 60 dBSL (A). Corresponding spectrograms (B) show elicited activity in the F0 bandwidth. Click responses of the same subject at the 3 corresponding intensities are shown in C.

medial geniculate body [67]. Neurophysiological spectral information regarding encoding was clearly identified at high intensities (50 and 60 dB SPL) but partially blurred at 40 dB SPL (Figure 1B). FFR mimicked the temporal features of the /ba/ stimulus with a 15.9 ms (± 1.3) interval at 60 dB SPL that decreased as intensity increased ((16.6 ± 1.5) at 50 dB SPL; (17.5 ± 1.9) at 40 dB SPL (Figure 2A). OR and click wave V latencies shortened as stimulation intensity increased (Figures 1A-

1C, 2B and 2C). At intensities lower than 40 dB SPL, the OR and FFR components were not clearly identified and poorly reproducible. Unlike the wave V evidenced until 0 dB SPL, consistent with behavioral hearing testing data, the OR and FFR require intensities above 40 dB SPL. Given that brainstem responses elicited by a 30 dB SPL /ba/ stimulus were not reproducible across participants, only components recorded at 40, 50 and 60 dB SPL were further evaluated.

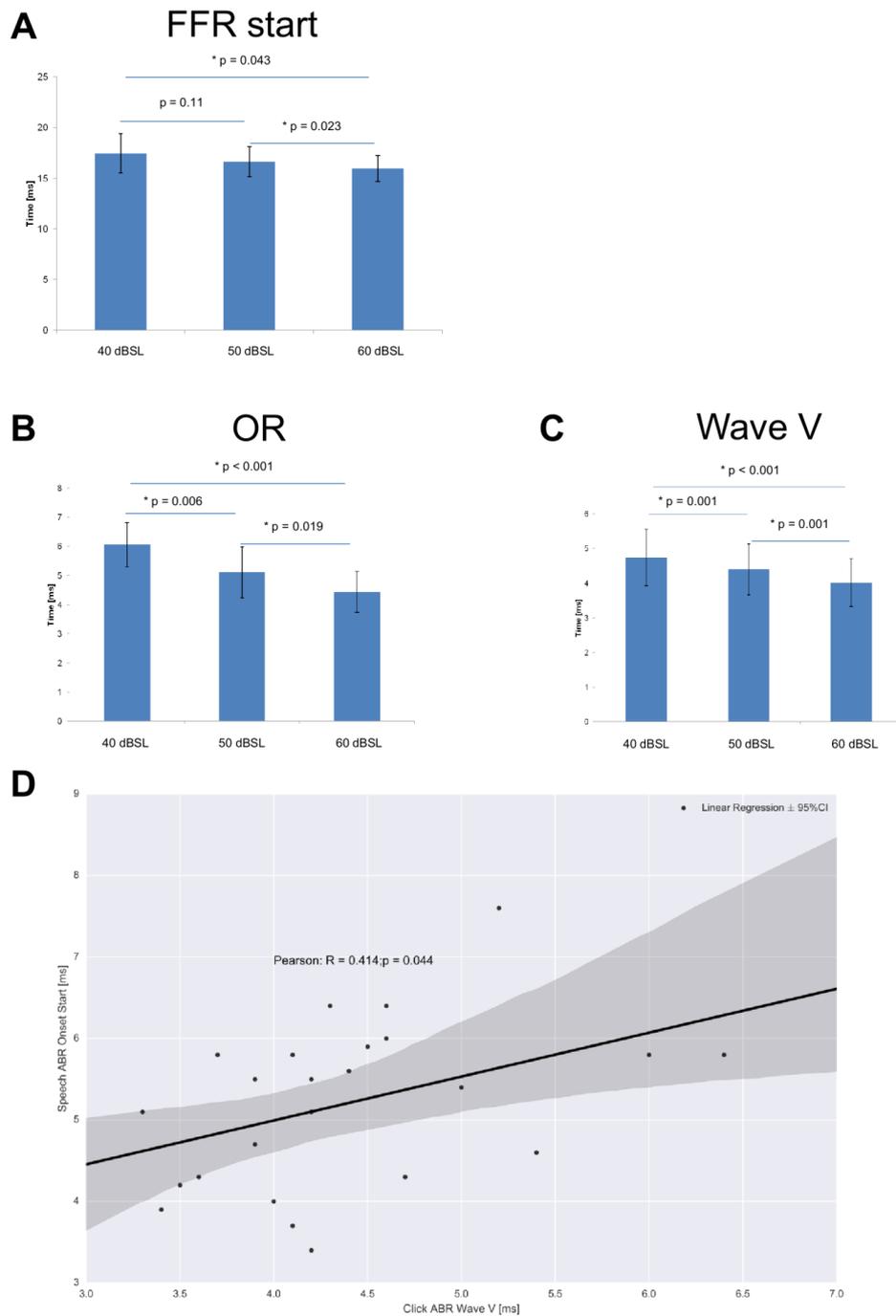


Figure 2: Effect of intensity on recorded latencies of OR and FFR components of the speech-elicited ABR and of wave V of the click-evoked ABR. Bar plot representations of the stimulation intensity effect on FFR (A, SD=1.92 at 40 dBSL, SD=1.5 at 50 dBSL and SD=1.30 at 60 dBSL), OR (B, SD=0.77 at 40 dBSL, SD=0.87 at 50 dBSL, SD=0.61 at 60 dBSL) and wave V latencies (C, SD=0.82 at 40 dBSL, SD=0.74 at 50 dBSL and 0.68 at 60 dBSL). Pairwise t-test values are provided. Linear regression revealed direct linkage between wave V latency and OR latency evolution patterns (R=0.414; p=0.044). Also shown is the 95% confidence interval of the regression.

Subcortical linear relation between waves V and OR as a function of intensity

Speech ABR and click ABR component latencies varied by intensity. Latency of the FFR start tended to become shorter relative to stimulation intensity increase (mastoid referenced RMANOVA $F(2, 14)=6.84$; $p<0.01$; average referenced RMANOVA $F(2, 14)=6.45$; $p=0.01$), with a significant decrease in latency from 50 to 60 dB SPL (mastoid referenced $p=0.007$, Figure 2A; average referenced $p=0.01$). Similar results were found regarding the FFR start whether mastoid or average referenced, with a significant decrease in latency from 60 to 40 dB SPL (average referenced $p=0.02$) but not between 40 and 50 dB SPL (average referenced

$p=0.41$). A significant effect of stimulation intensity on both latency of onset and wave V latency was found overall and between each of the three intensity conditions (onset, mastoid referenced $F(2, 14)=30.09$; $p<0.01$, all pairwise $p<0.02$, average referenced $F(2, 14)=15.92$, $p<0.01$, all pairwise $p<0.03$); wave V, $F(2, 14)=51.27$; $p<0.01$) (Figures 2B and 2C). The onset duration (waves V-A of the speech ABR) and slope did not vary as a function of stimulation intensity but its amplitude showed a tendency to decrease with decreasing intensity ($p=0.08$). Although FFR component amplitudes tended to decrease as the intensity decreased (Figure 1A), this relationship was not statistically significant. As shown in Figure 2D, latencies of OR and wave V across intensities correlated ($R(23)=0.414$; $p=0.044$).

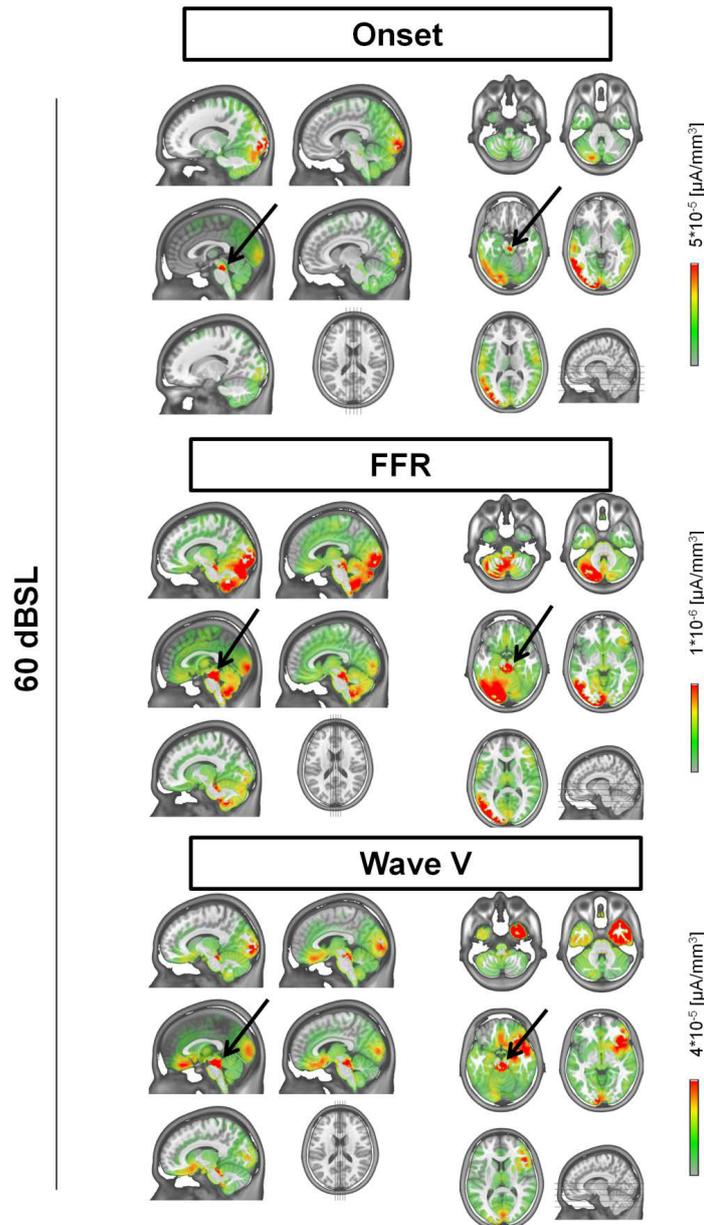


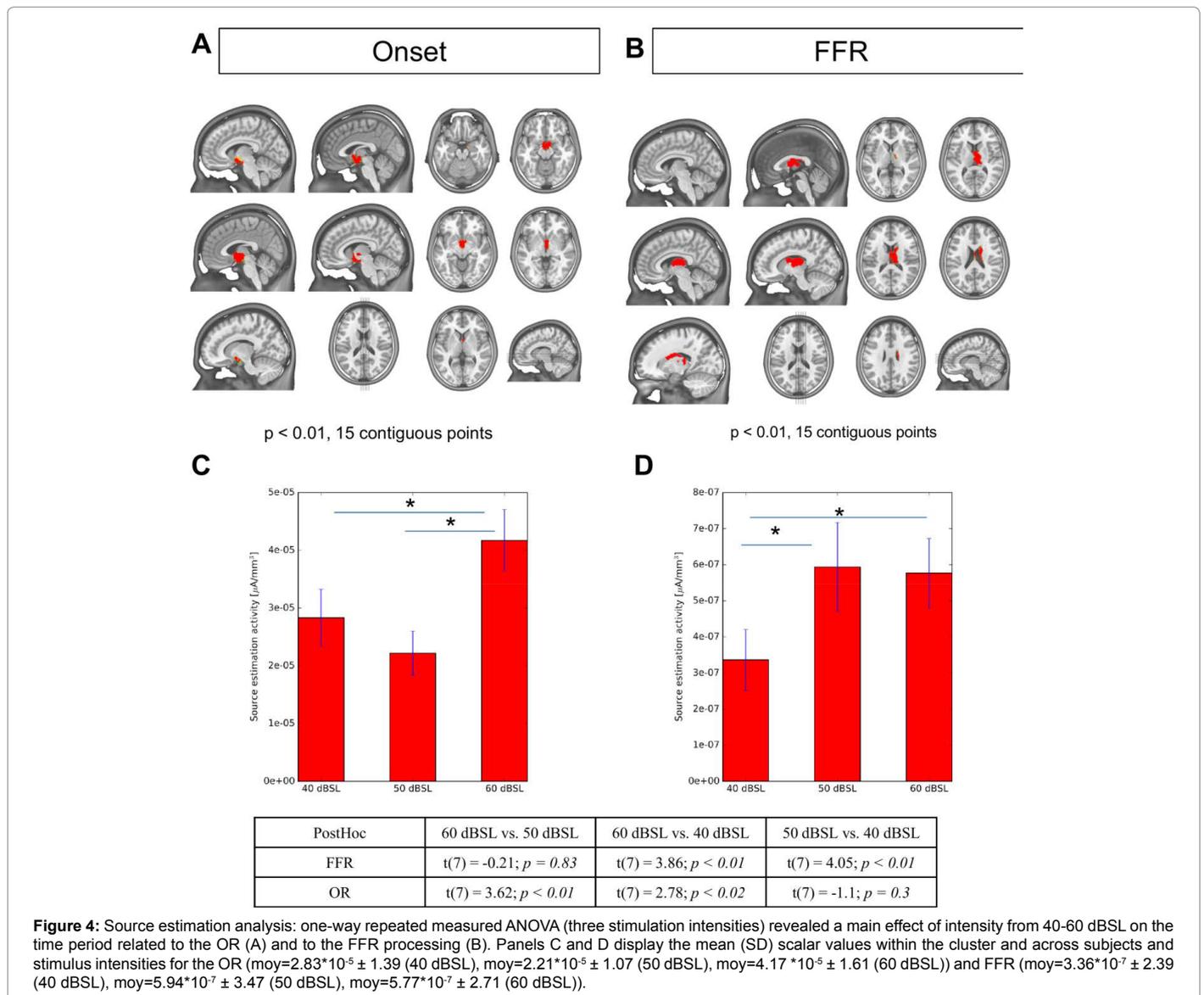
Figure 3: Identification of generator source performed using a distributed linear inverse solution (ELECTRA) applying the local autoregressive average (LAURA) regularization approach to address the non-uniqueness of the inverse problem. Mean of each individual's baseline-corrected LAURA source imaging are presented with generators of the OR (black arrow, upper panel), FFR (black arrow; middle panel) and wave V (black arrow; lower panel) are presented in sagittal (left) and axial (right) views. All views depict mean average activity in response to a 60 dBSL stimulus (either click or /ba/) while subjects were watching a silent movie (inducing the occipital activation seen on the different views).

Brainstem source location analysis of the neurophysiological mechanisms involved in click and speech processing

Spatiotemporal mapping of brainstem auditory responses: Brainstem click and speech ABR were characterized at 40, 50 and 60 dB SPL. Three time periods of activity were distinguished, corresponding to the time range of the wave V of the click-evoked response, the speech-evoked FFR and OR components. Despite some inter-subject variability, source localization methods revealed a common spatiotemporal pattern of activities involving the upper brainstem (midbrain) in the inferior colliculus area (Figure 3). At 60 dB SPL, during the time frame of the wave V, we observed a progressive activation of generator sources in the dorsal part of the upper brainstem in the inferior colliculus with a similar strength of brain activity as the onset response. However, the strength of the activity related to the subcortical source generators involved in FFR processing was 50-fold lower than the OR and wave V. Even though the source generators of the OR were mostly identified in the dorsal upper brainstem, some activity was revealed in the ventral part of the midbrain (Figure 3, upper panel). Activity related to the FFR

was predominantly localized to the caudal part of the upper brainstem. Although inverse solution did not differentiate the location of the OR and FFR in the dorsal part of the upper brainstem, sources involved in FFR processing exhibited lower subcortical activity intensity compared to the OR (Figure 3).

Brainstem source generators involved in OR and FFR processing show different sensitivity to stimulus level: Source estimation analysis one-way repeated measures ANOVA (condition=three stimulation intensities) revealed a main effect of intensity related to the onset processing in the hypothalamus (Figure 4A). FFR processing was more sensitive to stimulus intensity in the thalamic area, predominantly in the right thalamus (Figure 4B). Post hoc analyses of the effect per intensity revealed distinct activity patterns between OR and FFR processing (Figures 4C and 4D). While brainstem sources involved in OR processing showed sensitivity to stimulus intensity variation between 50 and 60 dB SPL, the activity related to the FFR generator sources varied greatly at lower intensities, between 40 and 50 dB SPL (Figure 4E).

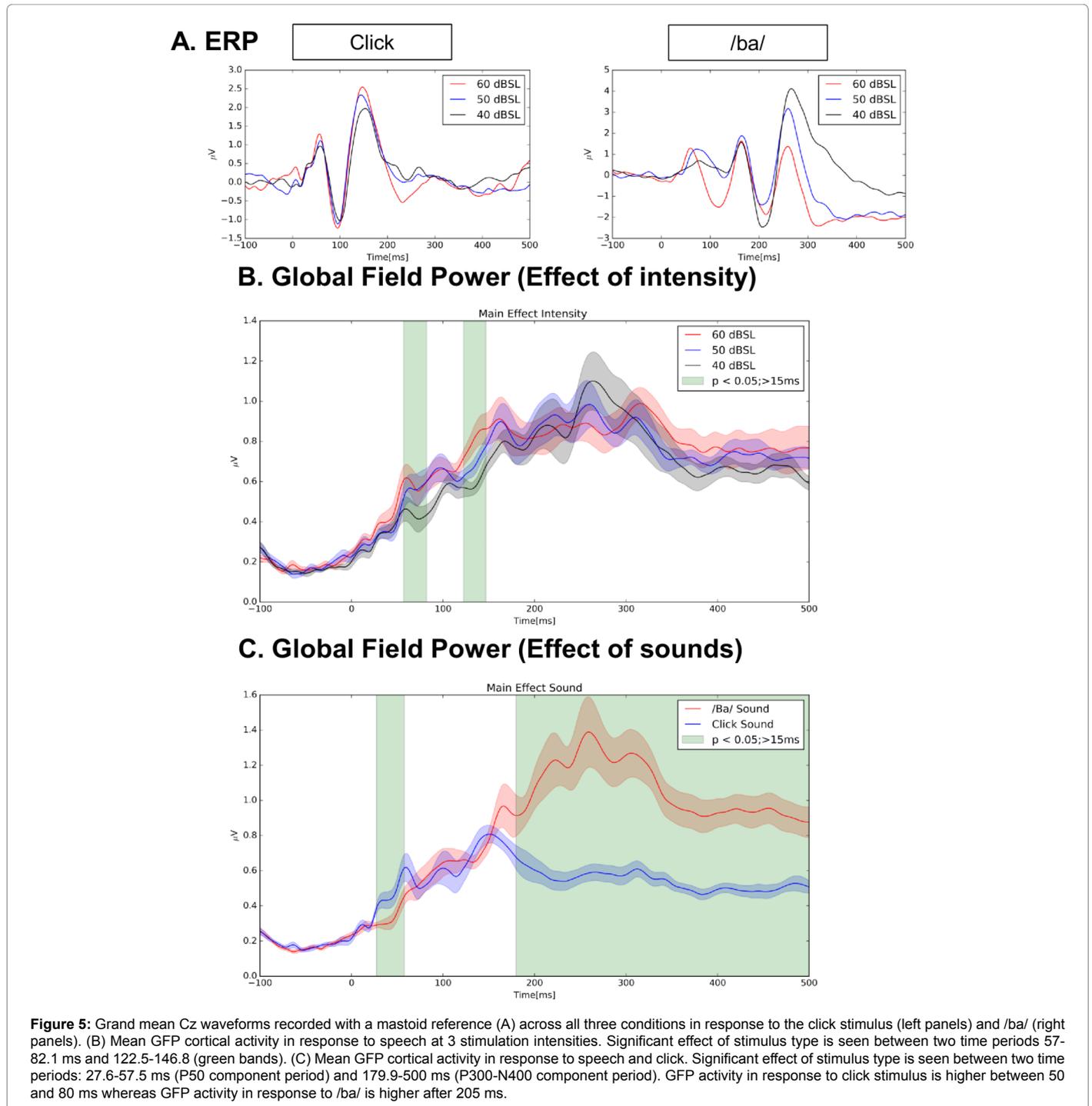


Cortical click-elicited response versus speech-evoked potential: Suggested patterns of processing

Cortical response peaks P50, N100, N200, P200, P300 and N400 for each stimulus and intensity (/ba/, click at 40, 50, and 60 dB SPL) are presented in Figure 5A. The 3×2 time-wise RMANOVA ($p < 0.05$, > 15 ms) at GFP level showed a main effect of intensity over 57-82.1 ms and 122.5-146.8 ms (Figure 5B) and a main effect of sounds over 27.6-57.5 ms (P50 component period) and 179.9-500 ms (P300-N400 component period) post-stimulus interval (Figure 5C). Visual inspection of

significant periods for the main effect of intensity revealed a difference between 40 dB SPL and the other intensities only for the first period (Figure 5B). The second period showed an association between intensity and GFP values ($60 > 50 > 40$ dB SPL). Visual inspection of relevant periods for the main effect of sound (Figure 5C) showed bigger GFP for clicks versus /ba/ for the P50 component period, but the P300-N400 component shows bigger GFP for /ba/ than for clicks.

The 3×2 time course analysis of variance (TANOVA) analysis provided similar results regarding the main effect of sound related to

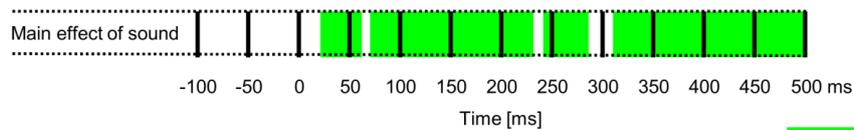


GFP. In addition, consistency of statistical topography maps across subjects and for each experimental condition showed the influence of sound (click and /ba/) on the main effect of sound periods (Figure 6B). Indeed, results provided evidence for a consistent pattern of active sources related to the P50 component period was consistent across all intensities for the click sounds conditions. In contrast, the P300-N400 component period was consistent for the /ba/ sound conditions across all intensities (Figure 6B).

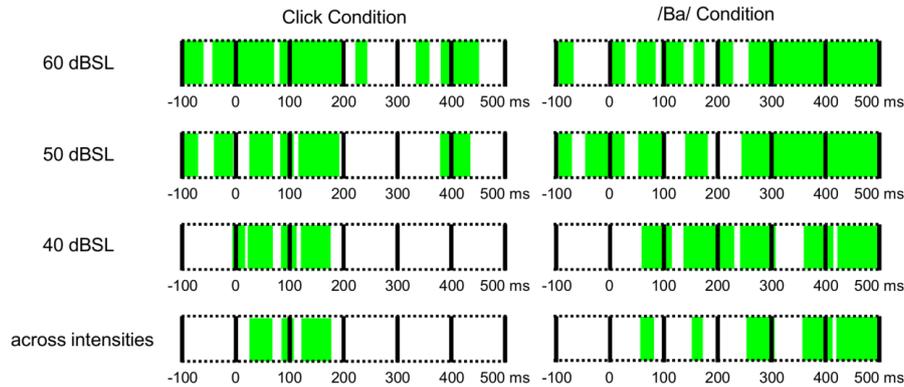
The GFP and topography main effect of sound (click and /ba/) for

the time periods of interest were used to define the source estimation parameters. The consistency analysis showed a clear relationship between the P50 component periods to the click sound condition, while the P300-N400 component period was related to the /ba/ sound condition. As expected, P50 component source estimation showed higher brain activity located in the right auditory cortex in response to the click (Figure 6C). In contrast, response to speech (/ba/) displayed maximum activity at P300-N400 located in the left auditory cortex, consistent with the well-known leftward specialization for linguistic function. Results showed the location of the generators involved in

A. TANOVA



B. Consistency Analysis



C. Location of source generators

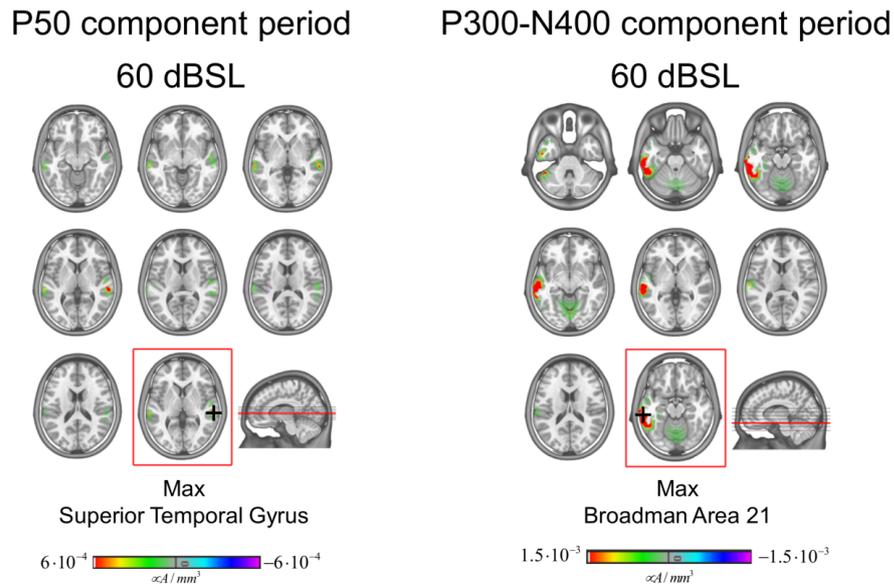


Figure 6: (A) Statistical analysis of topographical differences as a function of time (TANOVA analysis) revealed four time periods of statistically significant topographical differences between click and /ba/ (green bars). (B) Consistency analysis within each stimulus and condition (2 stimuli, 3 stimulation intensities) depicted the reproducibility of the topographical effect across participants as a function of time (significant periods: green bars). Analysis across intensities (lower row) shows a major effect around the P50 component period for the click and around the P300-N400 components period for the /ba/ condition. (C) Source location analysis over P50 and P300 periods (lower right panel) revealed greater activity in the right hemisphere in response to the click and in the left hemisphere to the /ba/.

click versus /ba/ processing under typical circumstances. Cortical areas activated during /ba/ versus click processing were statistically different (Figures 5C and 6A) and these results were consistent across all participants (Figure 6B).

Neurophysiological encoding at the brainstem level influences cortical processing

Given the aforementioned timeline and topography of speech and click processing in the cortex, a linear regression time-wise analysis was performed to evaluate the relationship between cortical GFP peak activity and brainstem encoding characteristics. There was a significant negative linear relationship between OR latency and GFP (cortical) over the 125.9-149.4 ms period (max at 139.6 ms; $p < 0.05$; > 15 ms) and a positive trend around 273.4 ms (Figure 7A). This demonstrated with a high temporal precision that an increase of OR latency induced a decreased activation related to the P100-N100 peak activity (139.6 ms; $R(23) = -0.596$; $p < 0.01$; Figure 7B). Similarly, P300-N400 peak activity tended to increase as the OR latency increased (273.4 ms; $R(23) = 0.354$; $p = 0.089$; Figure 7C). However, no effect on the OR slope, OR duration, FFR latency, FFR duration or FFR amplitude was found (all $p > 0.05$). We hypothesized that increases in wave V latency would correspond to statistically significant increases in cortical activity related to the P50 peak. All participants showed significant changes in P50 and P100 periods over a time period < 10 ms ($p < 0.05$).

Discussion

In this study, we investigated subcortical and cortical encoding of click and speech ABR at different stimulation intensities in healthy-hearing young adults. Findings suggested (1) a linear relationship between wave V and OR processing at the brainstem level that confirms shared mechanisms between the two components but with distinctive additional processing for the /ba/ stimulus, (2) location of the source generator of wave V, OR and FFR in the upper brainstem, (3) a robust effect of intensity in the thalamus and the upper brainstem for FFR and OR processing respectively, thereby providing direct evidence for differential processing of ORs and FFRs at the brainstem level, and (4) a positive temporal relationship between OR latencies at the brainstem level and cortical processing.

Wave V relationship to OR processing in the brainstem and comparison to previous studies of sound level on speech ABR

The present data regarding wave V and OR patterns suggest that while they share some underlying processing mechanisms (statistical correlation (Figure 2D) and a similar range of latencies), there is also a separate component distinctive to the processing of the /ba/.

Since the early report by Picton et al. [68] that described the OR by analogy to click ABR as the result of inferior colliculus cell activity, several reports have suggested similar generators between wave V and OR. However, the latencies of wave V and OR differ. The click ABR wave V typically occurs at 5.47 ms from stimulus onset for adults while mean OR latencies vary from 5 to 10 ms [69]. The existence of a distinctive mechanism is corroborated by animal studies [70,71] and modeling studies in well-hearing adult humans [72]. Similarly, previous works in learning-impaired children revealed typical click-evoked latencies while speech ABR latencies were disrupted [10], suggesting this process can be compromised in children with delayed speech ABRs [22]. Another facet is related to the structures of the stimulus itself: a speech stimulus such as the /ba/ syllable is a longer, complex stimulus compared to a click. Moreover, stop consonant identification relies on multiple acoustic cues conveying, for instance, voicing and manner of

articulation (both primarily conveyed in the temporal domain with possible contributions from spectral cues) and factors related to the place of articulation (encoded in the spectral domain [73-75] but with less high-frequency information than a click stimulus. A syllable evokes an OR (waves V-A) that is a transient complex component including responses to the onset of the sound, the onset of vocal cord vibration, and the offset of the sound [44]. This complexity was highlighted in this study by the lack of reproducibility of the OR below 40 dB SPL whereas wave V of the click ABR was reproducible down to 0 dB SPL. In addition to revealing, that high stimulation intensity is mandatory for precise analysis of speech ABR components, this also corroborates the concept that understanding language and speech requires higher intensities than the perception of sounds (clicks). An additional argument is made by Song et al. [22]: the encoding of click and speech auditory stimuli requires recruitment of different neural populations. The click-evoked brainstem potentials witness the integrity of the cochlea and ascending pathway while the speech ABR provides insight into the quality of the neural processing of complex sounds.

Source generators of the FFR, OR and wave V are located in the brainstem

The present data reveal the brainstem locations of the neural mechanisms responsible for different acoustic aspects of speech sound processing. Although it is generally accepted that the inferior colliculus houses click ABR wave V generators, the location of the FFR generators remains debatable. Electrophysiological recording of the latencies and spectral components of the FFR provide indirect evidence of a generator located in the brainstem [37,38,70,76-78]. Animal studies focusing on the discharge rate in the different nuclei strengthened the hypothesis of IC involvement [79]. A study using MEG suggested a contribution of the auditory cortex in the FFR consisting of cortical and subcortical components [39]. Other works support the concept of subcortical and cortical FFR generators [80,81]. However, evidence has also been reported for EEG activity to emerge from upper brainstem generators for the OR and FFR [82]. In Bidelman's report [72], the location, strength, and orientation of the generator source involved in speech-evoked FFRs were estimated using only a single pair of dipoles and thus may not have been able to dissociate multiple sources. Of note, it was previously suggested that dipole sources and source analysis should be further validated using co-recordings of MRI and functional EEG. The present data corroborate this hypothesis by suggesting generator sources of both OR and FFR are located bilaterally in the caudal part of the upper brainstem. Our source imaging results provide further evidence for human FFR sources arising bilaterally from the IC [34].

However, there are some limitations to our study. We used a passive listening set-up with a unique stimulus frequency to match the conditions relevant to clinical applications. However, the present stimulation intensity is above the F0 intensity generally used in the English literature [83,84]. Moreover, a 200 Hz stimulation is known to be above cortical neurons' phase-locking [85,86]. Together, these data and our source imaging results corroborate the idea of a stronger subcortical response to stimulus frequencies over 100 Hz [87]. As previously suggested [82], the relative involvement of the subcortical and cortical FFR sources may vary with stimulus frequency. Of practical implication when considering the clinical investigation of hearing impairment location, the FFR should be conceptualized as a spin-off from different generators whose involvement vary according to the stimulus characteristics [39,87,88]. Therefore, a 200 Hz F0 speech stimulus could be considered as a way to better separate subcortical from cortical FFR.

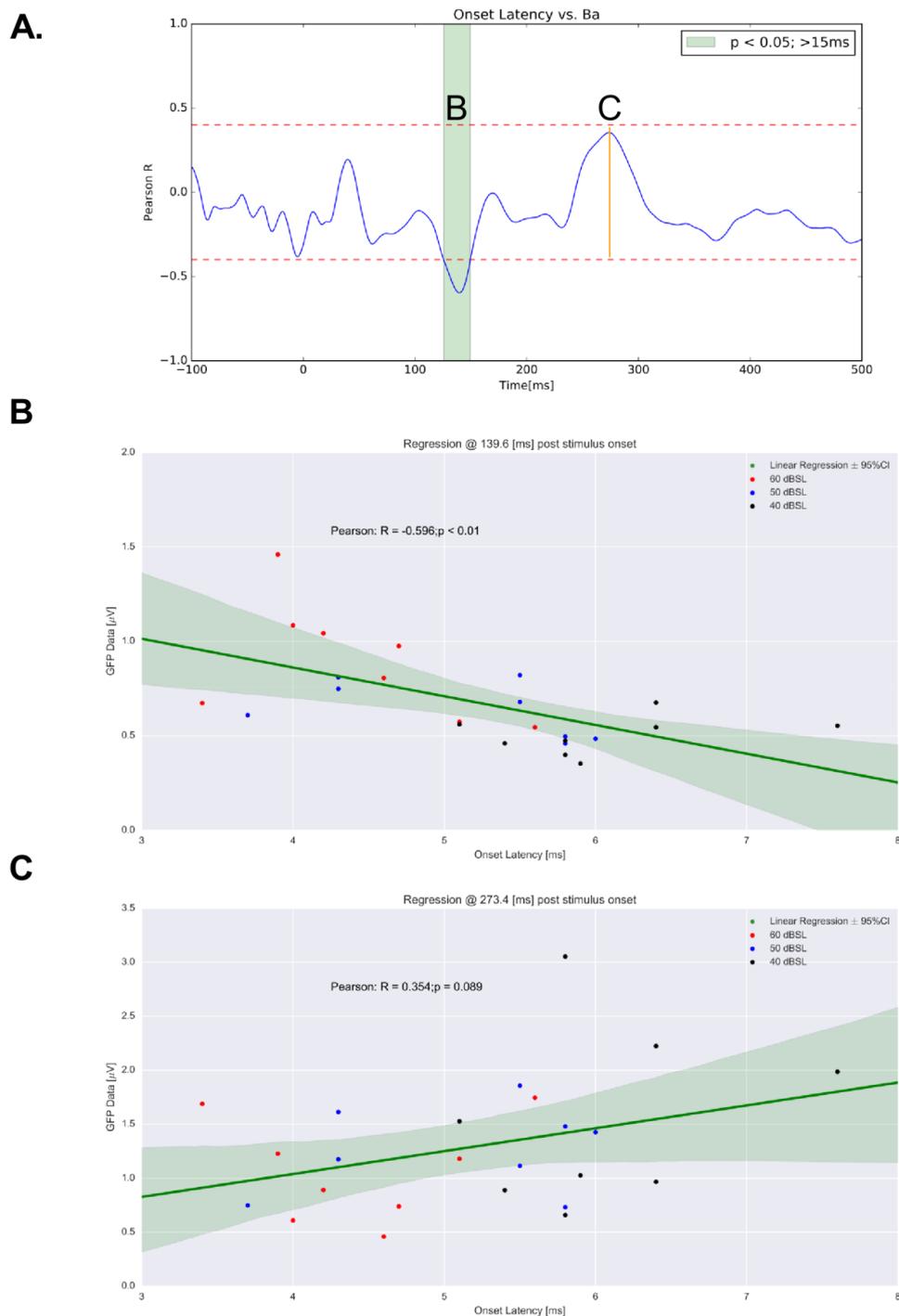


Figure 7: Estimation of the relationship between onset latency and cortical GFP in response to the /ba/ stimulus. (A) Pearson correlation between GFP at each time point and significance thresholds (dotted line). The latency on the OR showed one significant time period (>10 ms duration) between 125.9-149.4 ms and a trend around 273.4 ms. The linear regression analysis revealed a relationship between the latency of the OR at the subcortical level and the GFP activity (cortical) at 139.6 ms ($R(23)=-0.596$; $p<0.01$) (B), as well as at 273.4 ms ($R(23)=0.354$; $p=0.089$) (C).

OR and FFR processing involve different structures of the auditory system

Consistent with previous reports [13,38,44], OR and FFR latencies increased with decreasing intensities, suggesting a common mechanism between OR and FFR processing. Although OR and FFR

are thought to be processed in the same brainstem nuclei, the existence of distinct underlying mechanisms has been previously suggested [34,38,89-92]. We report for the first time results that suggest different sensitivities to intensity between OR and FFR at the brainstem level with the OR requiring generally higher levels to elicit sufficient neural synchrony for their generation. Furthermore, OR and FFR exhibited

direct relationships to two distinct areas (hypothalamus and thalamus respectively). Recent data showing the FFR to be different from the succession of wave V [82], as well as our present data, provide evidence that the onset and FFR components rely on different mechanisms at distinct portions of the auditory system, a finding directly visualized in the current study.

The thalamocortical network is known to be engaged in different brain functions including language, music and cognition [93]. Among its numerous roles, this network relays peripheral sensory signals to the primary sensory cortex [94] and carries information related to tone and rhythm through separate projection channels. Involvement of the middle geniculate body was previously suggested by fMRI studies revealing sound-related activation in the cochlear nucleus, superior olivary complex, inferior colliculus and medial geniculate body [95,96]. FFR integration of sound intensity in the MGB suggests pre-processing before transmitting information through projections to the auditory cortex for cortical emotional and cognitive appraisal [97,98]. In line with this observation, the medial geniculate body (MGB) may be a part of the plasticity of subcortical encoding [92]. The MGB has been implicated in the analysis of auditory communicative signals [99] as well as the processing of communicative signals loaded with emotions [100,101]. A proposed model for auditory communication [99] promotes the role of the inferior colliculus in decoding the spectral and temporal features of the signal, while the MGB is involved in the analysis of effect, highlighting the subcortical contribution as one of pre-processing before thorough cortical processing.

The mammillary bodies play a role in recognition memory (together with the anterior and dorsomedial thalamic nuclei) and spatial memory and learning and the hypothalamus acts as a control center for the autonomic nervous system.

Temporality of click and speech CAEP: comparison with previous observations

Clicks are predominantly encoded in the right hemisphere while speech is left lateralized. The present data are in congruence with the concept of a left hemispheric lateralization for vowel processing [102-104]. Our data support previous work suggesting there's a strong predisposition for speech sounds to be processed by the left hemisphere and non-speech signals by the right auditory cortex of the temporal lobe [105]. Subsequent mapping of phonology is more left lateralized [106]. Cortical GFP activity is mainly dependent around P50 for clicks while a speech stimulus influences the P300-N400 GFP-related activity. In the cortical area, we found greater activity strength in the P50 time frame for the click and at P300-N400 for the /ba/.

Subcortical encoding and cortical processing: disclosure of a direct temporal relationship

In the present study, we found a direct linkage between OR latency at the subcortical level and cortical GFP activity at two time points related to N100 and P300. Data gathered from MEG studies of evoked activity show possible phonological processing starting as early as 100 ms after sound onset [107]. However, cortical semantic and lexical processing begins between 200 and 300 ms after sound onset. Therefore, we speculate that an increase in OR latencies at the brainstem level induces a decrease in P100 GFP cortical activity while it increases the cognitive effort of semantic processing (P300). Previous reports suggest a relationship between OR latency at the subcortical level and higher incidence of language processing disorders, highlighting the influence of poor neural encoding at the brainstem level on higher cortical abilities

[90]. Other studies strengthen the hypothesis of a subcortical effect on cortical speech processing, for instance, reports related to hearing in adverse listening conditions [108,109] or related to optimize subcortical encoding in musicians [110,111]. The present data strengthen previous reports [110,111] by providing a direct temporal relationship between encoding by subcortical structures and cortical activity.

Comments on the methodology and importance of present findings

The 32-channel EEG system used in the present study provided a straightforward, rapid and non-invasive sensor application for subcortical and cortical auditory potential analyses but it does not permit the differentiation between one and multiple generators in the brainstem. Auditory scalp-recorded potentials reflect the engagement of multiple subcortical and cortical networks overlapping in both space and time. As such, it is difficult to ascribe intensity-related changes to a single neural generator. Different methods are based on varying assumptions related either to the geometrical, anatomical, or the electromagnetic properties of the brain that constrain the inverse problem. Nevertheless, processing techniques developed and previously used in our laboratory [42,112] have established their reliability and reproducibility. Aside from stimulus characteristics, the underlying principles of EEG and MEG may explain some differences observed in generator contribution between studies [39,83]. EEG fundamentals provide a better ability to establish accurate localization of neural generator compared to MEG [113,114]. Therefore, the EEG technique allows for direct visualization of gross changes in brainstem activation and functional involvement of the upper brainstem in the millisecond range while the MEG technique may jeopardize deep source signal extraction and interpretation [115] and thereby underestimate their contribution to FFR generation. Although MEG and EEG's distinct properties render the two modalities complementary in many respects [115,116], the operating costs of the MEG sensing technology still limit its implementation in clinical practice.

In the present report, we demonstrate that EEG is a reliable, affordable, practical, and straightforward modality applicable toward assessing the quality of speech encoding and identifying the neural generators that contribute to the scalp-recorded measures.

Conclusion

This exploratory study provides further information regarding the link between subcortical and cortical auditory circuitry. In addition, it showed the feasibility of a direct, noninvasive assessment of the location of subcortical generators involved in the processing of OR and FFR components. The approach described herein has great potential for enabling direct qualitative and topographical evaluation of auditory deficits and their mechanisms toward providing patients with optimized diagnoses and care strategies.

Funding

This work was supported by a FBM grant (Faculté de Biologie et de Médecine de Lausanne #29747).

Acknowledgement

We thank all the participants who volunteered their time. We are very grateful to Pr Clarke and Pr Murray for their help. Additionally, the authors thank Karen I. Berliner, Ph.D. for manuscript editing.

References

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

2. Picton TW (1999) Intracerebral sources of human auditory-evoked potentials. *Audiol Neurootol* 4: 64-79.
3. Jewett DL, Williston JS (1971) Auditory-evoked far fields averaged from the scalp of humans. *Brain* 99: 681-696.
4. Jewett DL, Romano MN, Williston JS (1970) Human auditory evoked potentials: Possible brain stem components detected on the scalp. *Science* 167: 1517-1518.
5. Dau T, Wegner O, Mellert V, Kollmeier B (2000) Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J Acoust Soc Am* 107: 1530-1540.
6. Elberling C, Don M (2010) A direct approach for the design of chirp stimuli used for the recording of auditory brainstem responses. *J Acoust Soc Am* 128: 2955-2964.
7. Picton TW, Skinner CR, Champagne SC, Kellett AJ, Maiste AC (1987) Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *J Acoust Soc Am* 82: 165-178.
8. Van Maanen A, Stapells DR (2009) Normal multiple auditory steady-state response thresholds to air-conducted stimuli in infants. *J Am Acad Audiol* 20: 196-207.
9. Ménard M, Gallego S, Truy E, Berger-Vachon C, Durrant JD, et al. (2004) Auditory steady-state response evaluation of auditory thresholds in cochlear implant patients. *Int J Audiol* 43: 39-43.
10. King C, Warrier MC, Hayes E, Kraus N (2002) Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neurosci Lett* 319: 111-115.
11. Banai JK, Hornickel E, Skoe T, Nicol S, Zecker N, et al. (2009) Reading and subcortical auditory function. *Cereb Cortex* 11: 2699-2707.
12. Krishnan A (2002) Human frequency-following responses: Representation of steady-state synthetic vowels. *Hear Res* 166: 192-201.
13. Russo N, Nicol T, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. *Clin Neurophysiol* 115: 2021-2030.
14. Anderson S, Parbery-Clark A, White-Schwoch T, Drehobl S, Kraus N (2013) Effects of hearing loss on the subcortical representation of speech cues. *J Acoust Soc Am* 133: 3030-3038.
15. Bidelman GM (2017) Communicating in challenging environments noise and reverberation, the frequency-following response: A window into human communication. *Nature* 61: 3-5.
16. Lee S, Bidelman GM (2017) Objective identification of simulated cochlear implant settings in normal-hearing listeners via auditory cortical evoked potentials. *Ear Hear* 38: e215-e226.
17. Kraus N, White-Schwoch T (2015) Unraveling the biology of auditory learning: A cognitive-sensorimotor-reward framework. *Trends Cogn Sci* 19: 642-654.
18. Johnson KL, Nicol TG, Kraus N (2005) Brain stem response to speech: A biological marker of auditory processing. *Ear Hear* 26: 424-434.
19. Skoe E, Kraus N (2010) Auditory brainstem response to complex sounds: A tutorial. *Ear Hear* 31: 302-324.
20. Tremblay KL, Friesen L, Martin BA, Wright R (2003) Test-retest reliability of cortical evoked potentials using naturally produced speech sounds. *Ear Hear* 24: 225-232.
21. Hornickel J, Knowles E, Kraus N (2012) Test-retest consistency of speech-evoked auditory brainstem responses in typically-developing children. *Hear Res* 284: 52-58.
22. Song JH, Banai K, Russo NM, Kraus N (2006) On the relationship between speech and non-speech-evoked auditory brainstem responses. *Audiol Neurootol* 11: 233-241.
23. Banai K, Abrams D, Kraus N (2007) Sensory-based learning disability: Insights from brainstem processing of speech sounds. *Int J Audiol* 46: 524-532.
24. Anderson S, Skoe E, Chandrasekaran B, Kraus N (2010) Neural timing is linked to speech perception in noise. *J Neurosci* 30: 4922-4926.
25. Rocha-Muniz CN, Befi-Lopes DM, Schochat E (2014) Sensitivity, specificity and efficiency of speech-evoked ABR. *Hear Res* 317: 15-22.
26. Kraus N, White-Schwoch T (2016) Neurobiology of Everyday Communication: What have we learned from music? *Neuroscientist*.
27. Elmer S, Hausheer M, Albrecht J, Kühnis J (2017) Human brainstem exhibits higher sensitivity and specificity than auditory-related cortex to short-term phonetic discrimination learning. *Sci Rep* 7: 7455.
28. Bidelman GM (2015) Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials. *J Neurosci Methods* 241: 94-100.
29. Bellier L, Veuillet E, Vesson JF, Bouchet P, Caclin A, et al. (2015) Thai-Van, speech auditory brainstem response through hearing aid stimulation. *Hear Res*, p: 325.
30. Gama N, Peretz I, Lehmann A (2017) Recording the human brainstem frequency-following-response in the free-field. *J Neurosci Methods* 280: 47-53.
31. Hoormann J, Falkenstein M, Hohnsbein J, Blanke L (1992) The human frequency-following response (FFR): Normal variability and relation to the click-evoked brainstem response. *Hear Res* 59: 179-188.
32. Galbraith GC (1994) Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalogr Clin Neurophysiol* 92: 321-330.
33. Møller AR, Jannetta PJ (1982) Evoked potentials from the inferior colliculus in man. *Electroencephalogr Clin Neurophysiol* 53: 612-620.
34. Sohmer H, Pratt H, Kinarti R (1977) Sources of frequency following responses (FFR) in man. *Electroencephalogr Clin Neurophysiol* 42: 656-664.
35. Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology* 47: 236-246.
36. Warrier CM, Abrams DA, Nicol TG, Kraus N (2011) Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hear Res* 282: 108-118.
37. Galbraith GC (2000) Putative measure of peripheral and brainstem frequency-following in humans. *Neurosci Lett* 292: 123-127.
38. Akhoun I (2008) The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clin Neurophysiol*. 119: 922-933.
39. Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ, et al. (2016) Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun* 7: 11070.
40. King A, Hopkins K, Plack CJ (2016) Differential group delay of the frequency following response measured vertically and horizontally. *J Assoc Res Otolaryngol* 17: 133-143.
41. Musacchia GD, Strait N, Kraus N (2008) Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear Res* 241: 34-42.
42. Murray MM, Brunet D, Michel CM (2008) Topographic ERP analyses: A step-by-step tutorial review. *Brain Topogr* 20: 249-264.
43. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9: 97-113.
44. Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: A tutorial. *Ear Hear* 31: 302-324.
45. Jäncke L, Mirzazade S, Shah NJ (1999) Attention modulates activity in the primary and the secondary auditory cortex: A functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 266: 125-128.
46. Parbery-Clark A, Skoe E, Kraus N (2009) Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci* 29: 14100-14107.
47. Beattie RC (1988) Interaction of click polarity, stimulus level and repetition rate on the auditory brainstem response. *Scand Audiol* 17: 99-109.
48. Cousineau G, Bidelman GM, Peretz I, Lehmann A (2015) On the relevance of natural stimuli for the study of brainstem correlates: The example of consonance perception. *PLoS One* 10: e0145439.
49. Werker JF, Tees RC (1984) Phonemic and phonetic factors in adult cross-language speech perception. *J Acoust Soc Am* 75: 1866-1878.
50. Dehaene-Lambertz G, Gliga T (2004) Common neural basis for phoneme processing in infants and adults. *J Cogn Neurosci* 16: 1375-1387.
51. Akhoun I (2008) Speech auditory brainstem response (speech ABR)

- characteristics depending on recording conditions, and hearing status: An experimental parametric study. *J Neurosci Methods* 175: 196-205.
52. Perrin F, Bertrand O, Pernier J (1987) Scalp current density mapping: Value and estimation from potential data. *IEEE Trans Biomed Eng* 34: 283-288.
53. Miodonska Z, Bugdol MD, Krecichwost M (2016) Dynamic time warping in phoneme modeling for fast pronunciation error detection. *Comput Biol Med* 69: 277-285.
54. Koenig T, Stein M, Grieder M, Kottlow M (2014) A tutorial on data-driven methods for statistically assessing ERP topographies. *Brain topogr* 27: 72-83.
55. Michel CM, Murray MM (2012) Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61: 371-385.
56. Khanna AA, Pascual-Leone C, Michel M, Farzan F (2015) Microstates in resting-state EEG: Current status and future directions. *Neurosci Biobehav Rev* 49: 105-113.
57. Koenig T (2002) Millisecond by millisecond, year by year: Normative EEG microstates and developmental stages. *Neuroimage* 16: 41-48.
58. Lehmann D (1971) Multichannel topography of human alpha EEG fields. *Electroencephalogr Clin Neurophysiol* 31: 439-449.
59. Lehmann D, Skrandies W (1980) Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr Clin Neurophysiol* 48: 609-621.
60. Skrandies W (1990) Global field power and topographic similarity. *Brain Topogr* 3: 137-141.
61. Guthrie D, Buchwald JS (1991) Significance testing of difference potentials. *Psychophysiology* 28: 240-244.
62. Koenig T, Kottlow M, Stein M, Melie-García L (2011) Ragu: A free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Comput Intell Neurosci* p: 938925.
63. Koenig T, Melie-García L (2010) A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topogr* 23: 233-242.
64. Spinelli L, Andino L, Lantz G, Seeck M, Michel CM (2000) Electromagnetic inverse solutions in anatomically constrained spherical head models. *Brain Topogr* 13: 115-125.
65. Gonzalez-Andino SL, Murray MM, Foxe JJ, de Peralta-Menendez RG (2005) How single-trial electrical neuroimaging contributes to multisensory research? *Exp Brain Res* 166: 298-304.
66. Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, et al. (2004) Grave de Peralta, EEG source imaging. *Clin Neurophysiol* 115: 2195-2222.
67. Galbraith GC, Bagasan B, Sulahian J (2001) Brainstem frequency-following response recorded from one vertical and three horizontal electrode derivations. *Percept Mot Skills* 92: 99-106.
68. Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials. I. Evaluation of components. *Electroencephalogr Clin Neurophysiol* 36: 179-190.
69. Richard C, Jeanvoine A, Veuillet E, Moulin A, Thai-Van H (2010) Exploration of the auditory system in humans: From click to speech auditory brainstem responses. *Neurophysiol Clin* 40: 267-279.
70. Parthasarathy A, Bartlett E (2012) Two-channel recording of auditory-evoked potentials to detect age-related deficits in temporal processing. *Hear Res* 289: 52-62.
71. Parthasarathy A, Datta J, Torres JAL, Hopkins C, Bartlett EL (2014) Age-related changes in the relationship between auditory brainstem responses and envelope-following responses. *J Assoc Res Otolaryngol* 15: 649-661.
72. Bidelman GM (2015) Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. *Hear Res* 323: 68-80.
73. Blumstein SE, Isaacs E, Mertus J (1982) The role of the gross spectral shape as a perceptual cue to place articulation in initial stop consonants. *J Acoust Soc Am* 72: 43-50.
74. van Tasell DJ, Greenfield DG, Logemann JJ, Nelson DA (1992) Temporal cues for consonant recognition: training, talker generalization and use in evaluation of cochlear implants. *J Acoust Soc Am* 92: 1247-1257.
75. Xu L, Thompson CS, Pfingst BE (2005) Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am* 117: 3255-3267.
76. Moushegian G, Rupert AL, Stillman RD (1973) Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalogr Clin Neurophysiol* 35: 665-667.
77. Kiren T, Aoyagi M, Furuse H, Koike Y (1994) An experimental study on the generator of amplitude-modulation following response. *Acta Otolaryngol* 511: 28-33.
78. Cunningham J, Nicol T, King C, Zecker SG, Kraus N (2002) Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hear Res* 169: 97-111.
79. Sinex DG, Henderson J, Li H, Chen GD (2002) Responses of chinchilla inferior colliculus neurons to amplitude-modulated tones with different envelopes. *J Assoc Res Otolaryngol* 3: 390-402.
80. Galambos R, Makeig S, Talmachoff PJ (1981) A 40 Hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci* 78: 2643-2647.
81. Kuwada S, Anderson JS, Batra R, Fitzpatrick DC, Teissier N, et al. (2002) Sources of the scalp-recorded amplitude-modulation following response. *J Am Acad Audiol* 13: 188-204.
82. Bidelman GM (2015) Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators and distinctions from the transient ABR. *Hear Res* 323: 68-80.
83. Bidelman GM, Yellamsetty A (2017) Noise and pitch interact during the cortical segregation of concurrent speech. *Hear Res* 351: 34-44.
84. Intartaglia B, White-Schwoch T, Kraus N, Schön D (2017) Music training enhances the automatic neural processing of foreign speech sounds. *Sci Rep* 7: 12631.
85. Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84: 541-577.
86. Abrams DA, Nicol T, White-Schwoch T, Zecker S, Kraus N (2017) Population responses in primary auditory cortex simultaneously represent the temporal envelope and periodicity features in natural speech. *Hear Res* 348: 31-43.
87. Tichko P, Skoe E (2017) Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. *Hear Res* 348: 1-15.
88. Zhang X, Gong Q (2017) Correlation between the frequency difference limen and an index based on principal component analysis of the frequency-following response of normal hearing listeners. *Hear Res* 344: 255-264.
89. Smith JC, Marsh JT, Brown WS (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol* 39: 465-472.
90. Wible B, Nicol T, Kraus N (2004) Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biol Psychol* 67: 299-317.
91. Purcell DW, John M, Schneider BA, Picton TW (2004) Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am* 116: 3581-3593.
92. Russo N, Nicol T, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. *Clin Neurophysiol* 115: 2021-2030.
93. De Witte L, Brouns R, Kavadias D, Engelborghs S, De Deyn PP, et al. (2011) Cognitive affective and behavioural disturbances following vascular thalamic lesions: A review. *Cortex* 47: 273-319.
94. McCormick DA, Bal T (1994) Sensory gating mechanisms of the thalamus. *Curr Opin Neurobiol* 4: 550-556.
95. Sigalovsky IS, Melcher JR (2006) Effects of sound level on fMRI activation in human brainstem, thalamic and cortical centers. *Hear Res* 215: 67-76.
96. Uppenkamp SM, Röhl M (2014) Human auditory neuroimaging of intensity and loudness. *Hear Res* 307: 65-73.
97. Huang CL, Winer JA (2000) Auditory thalamocortical projections in the cat: Laminar and areal patterns of input. *J Comp Neurol* 427: 302-331.
98. Doron NN, Ledoux JE (1999) Organization of projections to the lateral amygdala from auditory and visual areas of the thalamus in the rat. *J Comp Neurol* 412: 383-409.

99. Pannese A, Grandjean D, Frühholz S (2015) Subcortical processing in auditory communication. *Hear Res* 328: 67-77.
100. Cappe C, Morel A, Barone P, Rouiller EM (2009) The thalamocortical projection systems in primate: An anatomical support for multisensory and sensorimotor interplay. *Cereb Cortex* 19: 2025-2037.
101. Cappe C, Rouiller EM, Barone P (2009) Multisensory anatomical pathways. *Hear Res* 258: 28-36.
102. Vihla M, Salmelin R (2003) Hemispheric balance in processing attended and non-attended vowels and complex tones. *Brain Res Cogn Brain Res* 16: 167-173.
103. Parviainen T, Helenius P, Salmelin R (2005) Cortical differentiation of speech and non-speech sounds at 100 ms: Implications for dyslexia. *Cereb Cortex* 15: 1054-1063.
104. Gootjes L, Raji T, Salmelin R, Hari R (1999) Left-hemisphere dominance for processing of vowels: a whole-scalp neuromagnetic study. *Neuroreport* 10: 2987-2991.
105. Tervaniemi M, Hugdahl K (2003) Lateralization of auditory-cortex functions. *Brain Res Brain Res* 43: 231-246.
106. Turkeltaub PE, Coslett HB (2010) Localization of sublexical speech perception components. *Brain Lang* 114: 1-15.
107. Shtyrov Y, Kimppa L, Pulvermüller F, Kujala T (2011) Event-related potentials reflecting the frequency of unattended spoken words: A neuronal index of connection strength in lexical memory circuits? *Neuroimage* 55: 658-668.
108. Parbery-Clark A, Skoe E, Kraus N (2009) Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci* 29: 14100-14107.
109. Anderson S, Parbery-Clark A, Yi HG, Kraus N (2011) A neural basis of speech-in-noise perception in older adults. *Ear Hear* 32: 750-757.
110. Musacchia G, Strait D, Kraus N (2008) Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear Res* 241: 34-42.
111. Bidelman GM, Weiss MW, Moreno S, Alain C (2014) Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur J Neurosci* 40: 2662-2673.
112. Knebel JF, Javitt DC, Murray MM (2011) Impaired early visual response modulations to spatial information in chronic schizophrenia. *Psychiatry Res* 193: 168-176.
113. Hamalainen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography - Theory, instrumentation and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65: 413-498.
114. Darvas F, Pantazis D, Kucukaltun-Yildirim E, Leahy RM (2004) Mapping human brain function with MEG and EEG: Methods and validation. *Neuroimage* 23: 289-299.
115. Baillet S (2017) Magnetoencephalography for brain electrophysiology and imaging. *Nat Neurosci* 20: 327-339.
116. Baumgartner C (2004) Controversies in clinical neurophysiology. MEG is superior to EEG in the localization of interictal epileptiform activity. *Con Clin Neurophysiol* 115: 1010-1020.