# Deconstructing the Cortical Sources of Frequency Following Responses to Speech: A Cross-species Approach

**Authors:** G. Nike Gnanateja[1], Kyle Rupp[2], Fernando Llanos[3], Madison Remick[2], Marianny Pernia[4,7], Srivatsun Sadagopan[4,5,7,8,*], , Tobias Teichert[4,5,6*], Taylor Abel[2, 5,*], and Bharath Chandrasekaran[1, 4*]

**Joint corresponding authors*

[1]Department of Communication Sciences and Disorders, University of Pittsburgh
[2]Department of Neurological Surgery, UPMC Children's Hospital of Pittsburgh
[3]Department of Linguistics, The University of Texas at Austin
[4]Center for Neuroscience, University of Pittsburgh
[5]Department of Bioengineering, University of Pittsburgh
[6]Department of Psychiatry, University of Pittsburgh
[7]Department of Neurobiology, University of Pittsburgh
[8]Center for the Neural Basis of Cognition, University of Pittsburgh

**Author Contributions:** G.N.G., B.C., T.J.A., T.T., and S.S. conceptualized the study. K.R., M.R., and T.J.A. collected the sEEG data. T.T collected the macaque data. M.P., and S.S. collected the guinea pig data. G.N.G., K.R., and B.C. analysed the sEEG data. G.N.G., T.T., S.S., and B.C. analysed the macaque and guinea pig data. F.L.L., G.N.G., and B.C performed the Representational Similarity Analysis. G.N.G., B.C., T.J.A., T.T., S.S., K.R., M.R., F.L.L., and M.P., wrote the manuscript.

**Abstract:**

Time-varying pitch is a vital cue in the processing of speech signals. Neural processing of time-varying pitch cues in speech has been extensively assayed using scalp-recorded frequency-following responses (FFRs), which are thought to reflect integrated phase-locked activity from neural ensembles exclusively along the subcortical auditory pathway. Emerging evidence however suggests that the auditory cortex contributes to the FFRs as well. However, the response properties and the relative cortical contribution to the scalp-recorded FFR are only beginning to be explored. Here we used direct intracortical recordings from human subjects and animal models (macaque, guinea pig) to deconstruct the cortical sources of FFRs and leveraged representational similarity analysis as a *translational bridge* to characterize similarities between the human and animal models. We found robust FFRs in the auditory cortex that emerged from the *thalamorecepient layers* of the auditory cortex and contributed to the scalp-recorded FFRs via volume conduction.

## INTRODUCTION

Time-varying pitch patterns are a vital component of all spoken languages. Periodicity, a critical cue for time-varying pitch(1–3) can be non-invasively assayed using the scalp-recorded frequency following responses (FFRs) in animal(4–6) and human models(1,6–14,14,15). Scalp-recorded FFRs reflect phase-locked activity from neural ensembles across the ascending auditory pathway(16–18) and provide an integrative and non-invasive snapshot of pitch encoding in neurotypical and clinical populations(1,6–14,14,15). In neurotypical populations, FFRs have been leveraged to demonstrate experience-dependent shaping of pitch patterns at pre-attentive stages of auditory processing. Periodicity encoding, as indexed by the FFRs, is found to be awry in neurodevelopmental disorders(13,19,20,20–25), acquired neurological disorders(26–28), and aging(29–34). Despite the critical theoretical contributions to our understanding of human auditory plasticity and the potential as an easy-to-record non-invasive biomarker, the neural sources of the scalp-recorded FFRs to pitch patterns are poorly understood. For more than three decades, the inferior colliculus and the cochlear nucleus were considered the primary neural sources of the FFRs(5,10,11,35–37). Emerging research demonstrates a robust cortical contribution to the scalp-recorded FFRs(12,38–41). The neural dynamics and other response properties of the putative cortical ensembles contributing to the FFRs are poorly understood and will be the main focus of the current study.

Brainstem sources of scalp-recorded FFRs have been confirmed using cryogenic cooling of the IC(5) and resection of colliculo-cortical pathways(42) in animal models. Multichannel EEG recordings in humans also show that the subcortical sources predominantly contribute to the scalp-recorded FFRs(43–46). Furthermore, prior work has also suggested that temporal periodicity code of pitch is transformed into a rate or rate-place code in the upper brainstem(1). Recent studies

using magnetoencephalography (MEG) have challenged these accounts and shown strong cortical contributions to the scalp FFRs, with a distinct rightward cortical asymmetry(38–40). However, MEG studies require inferences based on distributional source modelling approaches which are relatively less sensitive to deep brain sources. But FFRs to speech stimuli have also been demonstrated in the auditory cortex using direct intracortical recordings(47,48) . However, the detailed arrangement of the current sources and sinks that give rise to these cortical FFRs are only beginning to be understood. As such it remains unclear how far the cortical FFRs are volume-conducted and contribute to the scalp-recorded FFRs.

To this end, we used an integrative cross-species (human, macaque, guinea pig (GP)) and cross-level approach to deconstruct the cortical contribution to the FFR with unprecedented mechanistic detail. We harmonized EEG recording procedures and elicited robust FFRs to time-varying pitch patterns across species. Cross-species translation is facilitated by using identical (and unfamiliar) speech stimuli (Mandarin tone patterns) that have extensively been used in prior work to elicit the FFR(49–54). Representational similarity analysis (RSA) served as a translational bridge(55,56), allowing us to quantify similarities in pitch encoding across species despite inherent fundamental anatomical differences and arousal states. We used laminar depth-electrodes in macaques and GPs to examine the layer-specific source of the cortical FFRs. In the animal models, cross-level translation was facilitated by acquiring EEG and intracranial neural recording methodologies in the same animal.

We demonstrate that robust FFRs to the Mandarin tone stimuli can be elicited from the auditory cortex in the human, macaque and GP models. RSA showed high representational similarity between the cortical FFRs in all three species. We further demonstrate, using layer-specific FFR recordings, that neurons in the thalamorecipient layers of the primary auditory cortex

are the primary sources of cortical FFRs. Further, FFRs recorded from the cortex shared strong power coherence with the scalp-recorded FFRs at low fundamental frequencies. Finally, we disentangled the relative subcortical and cortical contributions to the volume-conducted scalp FFRs by leveraging blind source separation in the same animal. Taken together, our results point to robust cortical FFRs to time-varying pitch patterns across species that contribute to the scalp-recorded FFRs.

## RESULTS

### FFRs in the human auditory cortex to vowels with time-varying pitch contours

Stereoelectroencephalography (sEEG) was performed in a nine-year-old male who required invasive monitoring(57) for drug-resistant seizures. Twenty electrode trajectories were planned in the participant, six of which included the supra-temporal plane of each hemisphere to investigate the insulo-opercular cortex. All the electrodes were stereotactically implanted, (Fig. 1A) as previously described. Such an opportunity to simultaneously record from bilateral temporal lobes is rare and paved the way to systematically study the properties of FFRs to speech recorded from the bilateral temporal lobes in humans. The participant listened to a range of speech and non-speech stimuli used to evaluate the FFRs recorded from the immediate proximity of auditory cortex. We analyzed time- and phase-locked neural activity to the periodicities in the stimulus. Robust FFRs (Fig 1C & 2A) with amplitudes above pre-stimulus baseline ($p<0.05$ on permutations-based t-tests between pre-stimulus baseline and FFRs on bootstrapped FFR trials) were observed in the electrode trajectories along the superior temporal gyrus (STG), planum temporale (PT), and the Heschl's gyrus (HG). The other two electrode trajectories in the planum polare and planum temporale did not show any FFR-like responses bilaterally. Thus, FFR analyses were restricted to the electrode trajectory bilaterally along the superior temporal plane.

We used four-pitch variants of the vowel /yi/ (Fig 1B), referred to as Mandarin tones to elicit the FFRs (Fig 1C). Mandarin tones have been extensively used to record FFRs to study the neurophysiology of pitch processing and associated plasticity in humans (50,52,53,58). The Mandarin tones are phonetically described as; T1 (high-level F0), T2 (low-rising F0), T3 (low-dipping F0), and T4 (high-falling F0). Morphologically, the time-locked averaged sEEG responses to the Mandarin tones showed robust onset responses followed by FFRs that lasted throughout the stimulus duration. We refer to the FFRs recorded from electrode contacts in close proximity to or directly within the auditory cortex as 'cortical FFRs (cFFRs)' from here on. As is the case for scalp-recorded FFRs, the cFFRs closely followed the Mandarin tones' fundamental frequency (Fig 1B & D and 2B). All four Mandarin tones elicited robust cFFRs (Fig 1C & D) in the electrode trajectory that was inserted along the HG, PT, and STG. The cFFRs that showed the highest amplitudes and signal-to-noise ratios were derived from electrodes located in the HG bilaterally (Fig 2A) ($p<0.05$, Permutation based ANOVA followed by post-hoc paired t-tests on bootstrapped trials). Electrodes in the PT did not show any significant FFR amplitude above the pre-stimulus baseline ($p>0.05$, permutation-based t-tests on bootstrapped FFR trials)**.**

The magnitudes of the cFFRs were highest for tone stimuli with lower-F0, i.e., most robust in T3 (89-111 Hz) and T2 ( 109-133 Hz), followed by T1 (~129 Hz) and T4 (140 to 92 Hz) (Fig 1D). This pattern is clearly visualized within the cFFRs to T2 and T4, where strong phase-locking can only be observed when the F0 of the vowel is low and phase-locking declines when the F0 is high (Fig 1D).

## Cortical FFR latencies in human sEEG are not consistent with volume-conducted activity from the brainstem

We assessed the latencies of the cFFRs for vowel T3 as they were the strongest and present throughout the stimulus duration. The latencies of the onset portion of the cFFRs were 14-16 ms, which is much later than the latencies expected of brainstem responses. Similar to the onset latencies, cFFR latencies (based on cross-correlation lags with maximum correlation coefficient) in the right HG(59) were ~17 ms, while they were ~ 23ms in the left HG. These latencies are not consistent with the earlier neural conduction delays expected of inferior colliculus activity and suggest that the recorded cFFRs reflect phase-locking of post-synaptic potentials in cortical neurons.

## Hemispheric asymmetry in cFFRs

cFFRs to Mandarin tones showed a distinct hemispheric asymmetry, consistent with a prior study using MEG(38). The electrodes in the right hemisphere showed higher amplitude cFFRs to the Mandarin tones ($p<0.01$, permutation-based t-tests on signal-to-noise ratios on bootstrapped cFFR samples). The rightward symmetry is also seen in the phase-locking value (PLV) spectrograms to the Mandarin tones. Higher phase-locking in the right hemisphere was also seen in an additional set of non-speech stimuli i.e., click trains with repetition rates in the human pitch range (Supplementary Fig 1). We used a Hidden Markov Model (HMM) to decode the mandarin tones from the cFFRs. The cFFRs from the right hemisphere tracked the stimulus pitch better than in the left hemisphere. Consequently, decoding accuracies were higher in the right hemisphere than the left hemisphere (Fig 2C). The pattern of tone decoding errors (‘confusions’) correlated significantly ($p<0.05$) between cFFRs from the right hemisphere and the scalp FFRs from a set of 20 subjects, but the same was not true for the cFFRs from the left

hemisphere and the scalp FFRs (Fig 2D)., Despite this difference, multidimensional scaling analysis revealed similar clustering of tone FFRs across the scalp, right HG, and left HG (Fig 2E).

**Representational similarity analysis (RSA) of cross-species and cross-level FFRs**

As in the humans, we recorded EEG in both model systems in order to establish scalp-derived FFRs as a translational bridge between the three species. Recordings in the animal models used the same Mandarin tone stimuli previously used for the human subjects. The scalp-recorded FFRs in both model species showed FFR activity above the pre-stimulus baseline ($SNR_{Maq}$ = 3.2, $SNR_{GP}$ = 6.9) and showed the expected phase-locking to the F0 of the stimuli (Fig 1G and 1J). Both species showed FFRs that correlated ($r_{Maq}$ = 0.45, $r_{GP}$ = 0.48) with the stimulus at latencies ($Lat_{Maq}$ = 3ms, $Lat_{GP}$ = 3.5ms) expected of early brainstem responses.

Furthermore, we recorded local field potentials (LFPs) from electrodes in PAC to compare against the intracranial LFP recordings in the human epilepsy patient. Similar to the scalp-recorded FFRs, the intracranial LFPs in both model species also yielded strong amplitude cFFRs above the pre-stimulus baseline ($SNR_{Maq}$ = 18.3, $SNR_{GP}$ = 5) and readily showed the expected phase-locking to the F0 of the stimuli (Fig 1G and 1J). The latencies of the cFFRs ($Lat_{Maq}$ = 11.6ms, $Lat_{GP}$ = 15.3ms), however, were longer than scalp-recorded FFRs in both species.

We used RSA to quantify similarities between humans and animal models across different recording levels (intracortical vs. scalp) (Fig 3). RSA was performed on confusion matrices constructed from FFRs recorded using harmonized stimuli (four Mandarin tones) across-species and levels. Human scalp data were derived from FFRs recorded in 20 participants

from a previously published study(50). In the macaque and GP subjects, scalp FFRs were recorded from cranial EEG electrodes surgically implanted in the skull. In the macaque, intracranial data were recorded from an electrode positioned immediately above layer 1 of the primary auditory cortex (PAC) from a chamber implanted over the frontal cortex. In the GP, intracranial data were recorded from an electrode contact estimated to be positioned in putative layer 4 of PAC. The pattern of phase-locking of FFR and cFFR in the animal models was similar to that seen in the human, with phase-locking declining rapidly with increasing stimulus F0 (Fig 1D, 1F, and 1J).

We decoded the Mandarin tone categories from scalp and intracranially recorded FFRs for all species using an HMM classifier(50,52). The HMM classifier performed at above chance levels (>0.25) across species and levels in identifying the correct pitch patterns from the FFRs (principal diagonal, Figure 3A). Human and animal FFR confusion matrices were strikingly similar at both the scalp (Fig 3B) and intracortical levels (Fig 3C) ($p<0.05$ on Pearson's correlation of confusion matrices without the principal diagonal) with stronger similarity seen for the intracortical cFFRs (Fig 2C). However, the scalp FFRs also yielded subtle species-specific differences (with greater similarity between human and GPs, relative to macaque). Further visualization of the FFR decoding matrices (diagonals removed) using multidimensional scaling (MDS) with Procrustes analysis showed similar clustering across the different species and levels (Fig 3D). These similarities appear to be predominantly driven by the higher confusion between the stimuli with higher F0 content (T1:High flat and T4:High-falling) and the least confusions for stimuli with lower F0 content (T2:low-rising and T3: low-dipping). Taken together the RSA reveals largely similar FFR representations for the four tones across species, with higher similarity observed in the intracranial cFFRs across species.

**High-density intracortical recordings in animal models reveal the laminar distribution of cFFRs**

Although intracortical recordings from human subjects provide high spatial and temporal resolution, they are still prone to contamination by volume-conducted fields from the brainstem and subcortical nuclei, and do not provide cortical layer-specific information. To overcome this limitation, we turned to laminar recordings from multi-contact electrodes traversing all layers of PAC approximately perpendicular to the cortical sheet (Fig 4) in the two animal model systems. These recordings allowed us to compute current source densities (CSD), which reflect post-synaptic currents and the corresponding passive return currents in the local cortical populations. Current sinks and sources are independent of volume-conducted potentials from the brainstem and the midbrain. The CSDs can be used to determine whether the post-synaptic currents in cortical populations are phase-locked to the stimulus, and if so, at which cortical depth and at which latency they arise. In addition, these laminar recordings also allowed us to assess the prevalence, latency and cortical depth of multi-units that phase-lock the F0 of the stimulus.

Fig 4D & H show CSDs of the low-pass filtered local field potentials (1 – 70 Hz) aligned to the stimulus onset. This analysis identified expected patterns of sources and sinks for both animals that were used identify the putative location of cortical layer IV, as well as supra- and infra-granular layers. We then computed the CSDs using the same filter setting used for the FFRs. Fig 4B & F show a 30 ms long snippet of the CSD FFRs from the sustained portion of tone 3 for both species. Note the presence of several currents that entrained to the F0 of the stimulus. We will refer to these currents as cortical frequency-following currents (cFFC). The most prominent cFFC was located in putative thalamo-recipient layers IV and deep layer III (as determined using the onset CSDs). A second somewhat weaker, opposite polarity cFFC was

identified in infra-granular layers. There was also an indication of a third and even weaker opposite polarity cFFC in supra-granular layers.

The cFFCs in both species showed a strong correlation with the stimulus at latencies of ~10-25 ms in the GP and 12-25 ms in the macaque (Fig 4C & G). These latencies are consistent with a cortical origin. Only one infragranular cFFC in macaque had a latency of 6.3 ms that seemed inconsistent with a cortical origin. It is likely that this particular cFFC does not exclusively reflect post-synaptic activity, but rather a very large-amplitude spike that was isolated at this electrode contact and was bleeding into the frequency range of the FFR. Given the short latencies, it is likely that the spike in question corresponded to a passing thalamocortical fiber, rather than an infragranular cortical neuron.

In both species, the strongest and most prominent cFFCs were recorded in granular layers, and most likely reflect active postsynaptic currents in response to F0-locked thalamic input at basal dendrites. It is less clear if the cFFCs in infra and supra-granular layers reflect active postsynaptic currents which might be indicative of the propagation of phase-locked responses to these layers or if they exclusively reflect passive return currents. In order for the phase-locked activity to spread beyond thalamo-recipient layers, not only the post-synaptic input currents but also the output, i.e., their firing rates, would have to be entrained to F0. We thus assessed the frequency following in the multi-unit activity (MUA), a measure of neural firing rate of small clusters of neurons in the immediate proximity of the electrode contact (cross-correlation between stimulus and MUA $r_{Maq}$ = 0.39, $\mathrm{Lat}_{\mathrm{Maq}}$ = 8.3ms, in the electrode with MUA above baseline $\mathrm{SNR}_{\mathrm{Maq}}$ = 1.625). The presence of FFRs in the MUA suggest that the thalamo-recipient layers can phase-lock to F0. This provides some initial support for the notion that the thalamo-recipient layers not only receive strong phase-locked input from the thalamocortical

fibers, but also propagate the FFRs to downstream cortical layers, albeit with less robust phase-locking strength.

**Relationship between scalp and cortical FFRs**

Because both intracranial and scalp FFR recordings were obtained in the same macaque and GP, we used the opportunity to examine the power and latency across frequencies of the scalp and cortical FFRs to infer similarities. The cortical FFRs were higher in amplitude than the scalp FFRs, presumably due to the proximity of the electrodes to the cortical sources. The comparison of the spectral characteristics of the scalp and cortical FFRs were thus made by normalizing the spectral estimates. Compared to the scalp FFRs, the cFFRs from the PAC were predominantly composed of low-frequency F0 energy relative to higher harmonics (Fig 1 and 5B). Fig 5A shows the FFRs to tone 3 (low-frequency dipping contour) recorded from the scalp and cortex in the macaque and the GP. Cortical cFFRs in both species showed longer latencies than the scalp-recorded FFRs (Fig 5A & B). While the phase-locking of cFFRs to Mandarin tones were higher in the PAC than at the scalp (Fig 1), the decline in phase-locking with increasing frequency was similar at both the PAC and the scalp. This can also be seen in the difference in normalized power spectral density between the scalp and cortical FFRs at the high frequencies when normalized based on maximum spectral amplitude (Fig 5B).

Cross-spectral power analysis revealed that scalp and cortical FFRs shared strong power coherence with the stimulus near the F0 (70-110 Hz), which declined rapidly at higher frequencies in the cortical FFRs relative to the scalp FFRs (Fig 5C). This trend was similar in both animal models. This pattern indicated that the FFRs recorded at the scalp and the cortex were similar in power spectral density at the low-frequency regions. Such a pattern can be caused either by a single common source or by more than one source with similar spectral properties but

different temporal properties. We thus estimated the cross-correlation strength and latency between the scalp and cortical FFRs across frequencies in the time-frequency domain. The maximum correlation between the scalp and cortical FFRs was seen at frequencies <120 Hz. However, at these frequencies, the cortical FFRs showed delays of ~7 ms in the macaque and ~10 ms in the GP in comparison to the scalp FFRs. This indicates the presence of temporally disparate but spectrally overlapping neural sources of the scalp-recorded FFRs. These delays can also be prominently seen in the latencies of the cFFRs, which are higher than conventionally obtained scalp-recorded FFRs (Fig 5C).

We further applied a blind source separation approach to disentangle the spectro-temporally overlapping components that contribute to the scalp-recorded FFR. We used independent component analysis (ICA) as the source separation approach (Maq - Fig 6 and GP-Fig 7). In the macaque, we submitted all the scalp (33 electrodes) and intracortical electrodes (96 electrodes spanning the superior temporal plane, prefrontal, and pre-motor cortex to ICA decomposition. We extracted 12 ICAs that explained 96% of the variance. Among these, we focused on four ICAs each of which explained >10% of the variance individually, and who as a group explained 75% of the variance (Fig 6). ICA2 was consistent with a volume-conducted generator from the regions distant to all intracranial electrodes (Fig 6A). This is apparent from the widely distributed ICA weights across electrodes. Furthermore, ICA2 had a latency of 3.4 ms and showed prominent power coherence with the stimulus at F0 as well as the higher harmonics (Fig 6C). In contrast, ICA1 had a longer latency (18.1 ms), responded strongest to the F0, and exhibited a steeper gradient between electrodes in the superior temporal plane and motor/premotor cortex. The power coherence with the stimulus also declined in frequency faster than in ICA1. These three findings are consistent with a generator in the primary auditory

cortex. This putative cortical ICA1 also propagated to the scalp and contributed to the scalp-recorded FFRs. Similarly, topographies and latencies of ICA4 (lat=0.5 ms suggested a cortical origin and propagated to the scalp electrodes. ICA3 and ICA4 in contrast to ICA1 show spatial weights that are opposite in polarity and hence possibly emerged from different cortical sources with different orientations.

In the GP, we submitted the 24 electrodes along the layers of the PAC, and 2 electrodes placed on the scalp to ICA. The scalp electrodes were placed on the vertex of the scalp (Cz : midpoint of the head along both sagittal and coronal axes) and on the temporal surface of the scalp close to the auditory cortex (T4). Six ICAs were extracted that explained 99% of the variance in the FFRs. Among these, the first four components each explained >10% of the variance in the FFR amounting to a total of 93.5% (Fig 6). The weights of ICA2 were not modulated along the laminar electrode layout consistent with volume-conducted activity from distant brain regions, in this case most likely the brainstem (Fig 7A & C). ICA2 had a latency of 4.1 ms and showed strong power coherence with the stimulus F0 as well as the harmonics (Fig 7D). The subcortical ICA2 contributed almost the entire variance in the scalp electrodes placed at Cz, thus suggesting negligent contribution from other sources such as cortex. ICA1 (lat=9.7ms) and ICA3 (lat=12.3ms) showed maximum spatial weights around putative cortical layers 4 and 2/3 and contributed maximally to the variance at the electrode T4 that was placed very close to the cortical surface of A1. In contrast, it did not contribute to the scalp electrode at Cz. While ICA4 (lat =15.1ms) was also consistent with a cortical source, it did not explain meaningful amounts of variance in the scalp electrodes. Taken together, these results show that in the GP, the scalp electrodes placed at the midline are dominated by the subcortical sources, while those placed on the temporal scalp locations are dominated by cortical sources.

Thus, the ICA results of both macaque and GPs suggest that the scalp-recorded FFRs contain weighted mixtures of both cortical and subcortical sources in the brain. The contributions of both these sources are dependent on the orientation of the net electrical dipole and the location of the scalp electrode.

## DISCUSSION

Research spanning three decades has richly characterized the properties of subcortical FFRs(5,12,15,16,60,61). However, significantly less is known about the cortical FFR. Our study establishes a cross-species (human, macaque, guinea pig) and cross-level (intra-cranial, scalp) platform to study cortical FFRs and their contribution to the scalp-recorded FFRs.

Our cross-species and cross-level approach revealed several novel results (a) All three species (human, macaque, GP) readily exhibited FFR-like responses to Mandarin tone stimuli in both scalp and cortical recordings. (b) Better encoding of lower frequencies was a characteristic of the cortical FFRs in all three species. Cortical FFRs were further characterized by longer latencies that were inconsistent with a volume-conducted source in the brainstem. (c) The bilateral FFR recordings from the Heschl's gyrus in humans showed more robust encoding and representation of higher fundamental frequencies in the right HG relative to left (d) RSA revealed striking similarities in the cortical representation of F0 contours, firmly establishing the two animal species as viable models to study the characteristics of cortical FFRs to human speech stimuli. (e) Laminar recordings from the macaque and GP auditory cortices demonstrated the existence of cortical frequency-following currents (cFFC) to human speech sounds in thalamo-recipient layers of PAC. (f) Using simultaneous EEG and large-scale intracranial recordings in the animal models we identified and quantified the putative contribution of the auditory cortex to the scalp-recorded FFR. Taken together, the main findings demonstrate the

utility of the different species, the different recording approaches and the synergy that arises when using identical stimuli across all of them. Below we highlight and expand upon the key findings within and across species.

**Cortical FFRs in humans show distinct low-frequency prominence with a right hemisphere bias**

Previous studies used distributional source modeling applied to EEG or MEG data to study the cortical source of FFRs (38,40,43). These methods are prone to (1) depth resolution errors and (2) challenges posed by the ill posed inverse problem of distributed source modelling. Here we circumvented the challenges of source modeling by directly recording from electrodes in the HG bilaterally, and confirmed the existence of the cortical FFRs across both hemispheres in the same participant. Two observations strongly argue against volume-conducted subcortical contributions to the cortical FFR. First, the cortical FFRs phase-locked only to the fundamental frequency, while subcortical FFRs phase-lock to speech harmonics as high as 950 Hz (35,62). Second, the latency of cortical FFRs (17-23 ms) was significantly longer than the latencies expected of subcortical regions. These latency differences make it highly unlikely that the cortical FFRs reflected a volume-conducted subcortical signal(63,64). We also observed inter-areal and inter-hemispheric differences that were inconsistent with volume conduction of subcortical potentials.

Bilaterally, the electrodes in the HG showed substantially stronger FFRs compared to those in the PT. The striking decline in the FFR magnitude in the PT electrodes indicates an FFR source in the HG and that the neurons in the PT do not phase-lock to the periodicity of the stimulus. Consistent with previously reported rightward bias in the cortical FFR activity(38,41,65), we found a distinct rightward asymmetry of FFR magnitudes. We found that

the decline in phase-locking with F0 was steeper in the left HG than the right HG, a pattern that was also observed for FFRs to non-speech click trains (Supplementary Fig 1).

While previous studies used speech sounds with a synthetic low F0 (~100 Hz) to assess the FFR cortical source(38,43), we used dynamically varying speech stimuli with a range of varying F0 (80 to 150 Hz) to record the FFRs. A previous multichannel EEG-based study examining FFRs to a low-frequency vowel (F0 = 120 to 80 Hz, 0.11 Hz/sec slope) suggested that the auditory cortex does not phase-lock strongly to speech F0s above 100 Hz(43). However, few studies(48) have systematically assessed phase locking in the bilateral auditory cortices to speech stimuli with F0s as high as 150 Hz. Our intracortical recordings demonstrate that the human auditory cortex can phase-lock to F0 as high as 150 Hz. A similar pattern was also observed in the FFRs to click trains with rapid decline at ~150 Hz but persisted at least till 200Hz (Supplementary Fig 1). The higher phase-locking limit of cortical FFRs in the current study possibly stems from the higher spatial precision of near- compared to far-field recordings.

Cortical FFRs showed phase-locking only to the fundamental frequency (Fig 5B). Further, Mandarin tones elicited FFRs only in the HG and not in the PT. This may indicate a possible transformation of a temporal pitch code into a place or a rate-place code in the auditory association cortex. Our observations of the preferential encoding of the dynamic pitch contours in the right HG has not been observed using direct cortical recordings before and may underlie right hemispheric asymmetry for processing melodic and prosodic features in (non-native) speech and music (65–68). Additionally, the rightward asymmetry was also seen for click trains, which suggests that the asymmetry is driven by stimulus acoustics and is unlikely to be speech-specific. Right hemispheric preference has been reported earlier for pitch processing using functional neuroimaging approaches(66,69,70). The rightward asymmetry that we observe in

FFRs directly recorded from the HG reinforces the idea of right hemispheric dominance for pitch encoding in speech. Notably, the participant in the study did not have any previous experience learning Mandarin. Thus, experience, another important factor driving the hemispheric laterality of pitch processing(71), likely did not account for the observed laterality in this case.

Experience-dependent changes of FFRs to Mandarin tone stimuli is widely considered proof of experience-dependent subcortical plasticity(50,51,72). Our finding of a cortical contribution to FFRs elicited by these very same stimuli provide an alternative explanation for the experience-dependent changes of the FFRs. As a result, subcortical plasticity models(49,73,74) that were proposed based on the FFRs will need to be revisited with a focus on the relative contribution of cortical and subcortical ensembles to experience-dependent plasticity.

**Human and animal models show comparable cortical frequency following responses to speech signals**

To facilitate comparison of FFR representation across levels and species and establish homologies we leveraged representational similarity analyses. We assessed the mandarin tone decoding accuracies based on the FFR pitch contours. These decoding accuracies are a measure of the discriminability of FFRs to different speech sounds that can be readily compared across species and across levels(56,75). The primary objective of the RSA approach was to establish a translational bridge between the human and animal FFRs.  This allowed us to further deep dive into the sources of the FFRs at an anatomical resolution that is currently unavailable in humans.

Strong similarities in the representational structure between the cortical and scalp FFRs were seen in both human and animal models. The similarity across the species suggests that the underlying signal transformation in the auditory systems of all three species are similar in the

sense that they prioritize the representation of the same auditory features. Macaque and GP may thus be well-suited to help answer important questions about the FFR, its origins, and relations with other higher-order processing. The extent of representational similarity across species was lower for scalp-recorded FFRs than intra-cortically recorded FFRs. This reduced similarity likely reflects a large variability in the dipole orientation of the cortical source between species that can lead to substantial differences in the relative contribution of cortical sources to the scalp FFR. The similarity pattern also indicates that intra-cortical FFRs across species are comparable and could be valuable in assessing fundamental aspects of auditory processing and experience-related modulation.

Consistent with the sEEG results in the human participant, PAC of the macaque and GP showed robust FFRs to the different pitch patterns. Multidimensional scaling analyses revealed a high degree of similarity in the representation of different tone patterns. In particular, in all species, the falling tone (T4) was represented less robustly (more confusion) than the rising tone (T2) (less confusion). This suggests a cross-species similarity in preference for stimuli with rising, relative to falling pitch patterns(76).

We also found substantial representational similarity between cortical and scalp-recorded FFRs in all three species. It is important to note that the absolute physical characteristics of the FFRs, such as the latency, amplitude, and the signal-to-noise ratio vary widely across the species due to the differences in the anatomical characteristics. The RSA approach provides a framework for comparing the representation of the pitch contours in the FFRs across the species without the confounds of the differences in absolute physical characteristics. Thus, the inference about similarity across species from the RSA should not be extrapolated to the physical characteristics of the FFRs listed above.

Furthermore, experiments using laminar depth electrodes revealed cortical frequency following currents (cFFC) in the putative layers III and IV of the primary auditory cortices of the macaque and GP. The latencies of these cFFCs closely matched the latencies of the cortical FFRs recorded using single-tipped electrodes in immediate proximity to A1. This finding strongly suggests that the cortical FFRs do not just reflect volume-conducted potentials from subcortical nuclei, but are also very likely generated by phase-locked post-synaptic currents in neurons of the primary auditory cortex. We emphasize that this finding is not at odds with the well-established existence of subcortical sources in scalp-recorded FFRs(6,43,60,77). In fact, a prerequisite for the generation of the FFRs at the cortex is the processing of F0 by the IC, which is then transmitted to the auditory cortex via the thalamus. Our results show that cortical ensembles are '*a'* source for the FFRs and not '*the'* source of the FFRs.

**Cortical FFRs emerge in the thalamo-recipient layers of the primary auditory cortex**

We consider FFRs as being of cortical origin if they arise from phase-locked post-synaptic currents in cortical neurons regardless of whether the post-synaptic currents are driven by thalamic or cortical input. Conversely, we define FFRs as having a subcortical origin if they arise from phase-locked post-synaptic currents of neurons located in subcortical nuclei. This working model is schematized in in Fig 8.

With the laminar depth electrodes, we confirmed the existence of cortical FFRs across layers of PAC. We found that currents in thalamo-recipient layers readily entrain to F0 input from the thalamocortical fibers. The outputs (MUA) of the thalamo-reciepient layers also entrained to stimulus F0.Further, the FFR currents show phase shifts across the different layers. The phase shift across the cortical layers is suggestive of the extracellular electric field's destructive interferences that may dampen the net equivalent current dipole at the cortex.

However, considering the lower strength of the cortical currents beyond the thalamo-recepient layer, it is parsimonious to assume that the thalamo-recipient layers primarily drive the FFRs recorded at the cortical surface in the human sEEG and in the macaque using sharp electrodes.

Earlier studies using MEG and EEG, as well as our sEEG results are suggestive of a cortical source for the scalp-recorded FFRs. With animal models we were able to precisely identify the multiple cortical sources of FFRs across the cortical laminae (See Figure 6 and 7). While FFR currents have been demonstrated in earlier work (78,79), these currents were assessed only till fundamental frequencies of 100 Hz. Here we show that cortical FFR currents show phase-locked activity as high as 150 Hz and follow the Mandarin tones' pitch contours.

FFR current sinks were also observed in the deep layer V. It is not clear if these currents are just passive return currents or active currents in which case they would indicate the propagation of phase-locked activity to layer V neurons which give rise to cortico-collicular feedback projections(80,81). Paradigms using stimulus probability effects have shown systematic effects on the FFRs, and have been considered as a putative marker cortico-collicular processing(19,31,82–84). The use of such paradigms in future studies using laminar recordings will help confirm whether layer V indeed underlies contextual-modulation of the FFR(19,31,82–85), and if the layer V responses show passive currents from the other cortical layers or if they reflect active currents in the layer V outputs.

**Cortical contributions to extracranially recorded FFRs**

The use of animal models enabled us to explore the similarities between the scalp and intracortically recorded FFRs in the same subjects. The power coherence between the scalp and cortical FFRs further shows that the scalp-recorded as well as cortical FFRs show strong

responses at the F0 of the stimulus. However, power coherence in the cortical FFRs declined for higher frequencies suggesting that the cortical sources do not strongly contribute to scalp FFRs at higher harmonics. Further, at the F0, there was a strong correlation between the scalp and cortical FFRs. However, these correlations were only seen when the scalp FFRs were delayed to match the latency of cortical FFRs, which implies that cortical FFRs showed longer latency than the scalp FFRs. The scalp and intra-cortically recorded FFRs showed strikingly similar decoding of mandarin tones in both the macaque and the GP, but they differed in neural timing. The scalp-recorded FFRs had shorter latencies than the cortical FFRs but showed similar tracking of the vowel pitch contours on the representational similarity analysis.

It is important to note that the multiple neural ensembles in the auditory cortex phase-lock to the F0 of the stimulus and form dipoles with varying orientations and strengths. These dipoles mix both constructively and destructively based on cortical folding and the direction of current propagation. The cortical dipoles also mix constructively and destructively with the volume-conducted electric fields from the subcortical nuclei. Such volume-conduction results in a net resultant electrical field at the scalp that does not directly reflect underlying neural activity. This problem is compounded in the case of the FFRs as they show strong correlation across time, which makes it difficult to distinguish activities across different neural ensembles as they correlate with each other at different time lags(12,43).

We used independent component analysis with electrodes placed on the scalp and the putative cortical sources of the FFRs to estimate the relative contribution of the cortical sources to the scalp FFR. We found a component in both species that reflected a volume-conducted source presumably from the subcortical regions that propagated uniformly to all cortical and scalp electrodes and showed earlier latencies. This independent component (IC) explained most

of the variance in the scalp electrodes We also found strong ICs that were localized to the auditory cortex but also projected to the scalp. These ICs likely reflect the bulk signal from the cortex that propagates to the scalp. Not all putative cortical ICs contributed to the scalp-recorded FFRs, presumably because they arose from neural ensembles that had closed dipole orientations that did not favor volume conduction to the scalp. It is important to note that the contribution of the cortical ICs to the scalp channels differed based on the location of the scalp electrode. The GP especially showed a very specific pattern of differential cortical contribution based on scalp electrode location. The midline electrode predominantly picked up the subcortical component, while the electrode above auditory cortex predominantly picked up the cortical ICA. Such a clear electrode montage difference was not seen in the macaque. This pattern likely stems from the species differences in the location and orientation of the auditory cortex. In the GP the auditory cortex is largely on the surface of the temporal lobe with cortical layers oriented medio-laterally unlike humans and macaques where the auditory cortex is largely buried in the lateral sulcus with cortical layers oriented dorso-ventrally(86,87). This further implies that the contribution of the cortical FFRs, though lower than the subcortical contribution, is present and dependent on the electrode montage used, the location of the auditory cortex and the folding patterns of the cortical surface.

## Conclusions

We demonstrate the existence of the cortical source of FFRs with direct intracranial recordings to time-varying pitch patterns. Using an integrated cross-species approach with the macaque and GP models, we show laminar profiles consistent with a cortical source of the FFRs that emerge from the thalamo-cortical recipient layers and not from the volume-conducted activity of subcortical neurons. We also show that while subcortical sources dominate the scalp-

recorded FFRs, cortical sources do contribute to the scalp-recorded FFRs. Our research paves the way for a wide array of studies to investigate this cortical FFR source's relevance in auditory perception and plasticity.

## METHODS

### FFRs to Mandarin tones.

#### *Stimuli*

Four pitch variants of vowel /yi/ were used to elicit the FFRs. These pitch patterns are meaningful linguistic pitch patterns in Mandarin and have been extensively used in studies that have used FFRs to study online modulation, plasticity, and learning (49,50,52–54,88). The Mandarin tone patterns are minimally distinguishable by the pattern of their fundamental frequency (F0) contour. The minimally contrastive F0 patterns are phonetically described as T1 (high-level, F0 = 129 Hz), T2 (low-rising, F0 ranging from 109 to 133 Hz), T3 (low-dipping, F0 ranging from 89 to 111 Hz), and T4 (high-falling, F0 ranging from 140 to 92 Hz). These stimuli were synthesized based on the F0 patterns derived from natural male speech production. All stimuli had a sampling rate of 48000 Hz and were 250 ms in duration. Stimuli were delivered using ER-3C insert earphones with the volume adjusted to a comfortable intensity level. The stimuli were presented in both condensation and rarefaction polarities to minimize potential contamination of the neural responses by the stimulus artifact and pre-neural cochlear microphonics(8). The stimuli were presented in a pseudorandom manner where each stimulus had a ¼ probability of occurrence. The overall number of stimulus sweeps per Mandarin tone differed across species, with at least 250 sweeps were obtained in each species; 800 (400 in each polarity for Human sEEG, 1000 (500 in each polarity) in human EEG, 1000 (500 in each polarity) in macaque, 250 (125 in each polarity) in the guinea pig intracortical recordings, and ~1658 (~829 in each polarity) in the guinea pig scalp recordings. The inter-stimulus interval for sEEG was 50 ms for human subjects, 250 ms for macaques, and 100 ms for GPs.

## FFRs to Click Trains

Click trains were used to assess the phase-locking limits of the cortical ensembles encoding the FFRs. These stimuli were used only in the human sEEG recordings. The use of these click trains were based on an earlier study investigating the coding of repetitive transients using intracortical recordings(89). Two sequences of repetitive click trains were used (89). The first sequence S1 consisted of 1s long click trains. The repetition rates in S1 were 4, 8, 16, 32, 64, and 128 Hz. The second click sequence S2 consisted of 0.16s long click trains. The repetition rates in S2 were 25, 50, 100, 125, 150, and 200 Hz.

Phase-locking across trials was assessed using the intertrial phase coherence of the local field potentials at the stimulus repetition rate. PLV spectrograms were extracted similar to the FFRs to vowels. The PLVs were averaged across time to obtain a PLV spectra. The PLV at the stimulus repetition rate was extracted at each electrode. The onset responses in the LFPs were too large and obscured the sustained responses to the click repetition rate. For better visualization of the sustained FFRs to the click repetition rate, the high-gamma power responses (Naplib Toolbox(90) ) are plotted.

## Intracranial and Scalp Electroencephalography in Humans

### *Participants*

*Human Participant for sEEG*

A nine-year-old boy with drug-resistant epilepsy participated in the study. The participant was right-handed, was a native speaker of English, and was attending grade 4 in school. The participant underwent stereoelectroencephalography (sEEG) monitoring of the bilateral temporal lobes for localization of his seizure foci. The choice of electrode insertion in the participant was

based purely on clinical necessity for evaluation of focal epilepsy. Such an opportunity to study record sEEG from both temporal lobes to study bilateral auditory processing in the same participant is rare . This child had no other relevant medical history. The participant's family gave written informed consent to participate in the study. All research protocols were approved by the Institutional Review Board of the University of Pittsburgh.

*Human Participants for EEG*

Data from a previously published study(50) with twenty participants in the age range of 18 to 24 years (12 females) was reanalyzed in this study. All the participants were monolingual native speakers of English. All the participants had hearing sensitivity within 20 dB HL across octave frequencies from 250 to 8000 Hz. Written informed consent was obtained from the participants before inclusion in the study. The research protocols used were approved by the Institutional Review Board of the University of Texas, Austin.

***Electrophysiological recordings in Humans***

*Stereoelectroencephalography in Humans*

Stereoelectroencephalography (sEEG) electrodes were stereotactically inserted in each hemisphere using robot-assisted implantation(57,91). Twenty trajectories were inserted along in different brain regions to test seizure localization hypotheses based on noninvasive evaluations(92). Each electrode had between 8-12 cylindrical contacts with a length of 2mm and a diameter of 0.8mm. The distance between each electrode contact was 3.5mm. Anatomical locations of the electrode sites were obtained using high-resolution CT and structural MRI. The electrode locations from the CT scan were coregistered with the structural MRI to precisely locate the anatomical locations of each electrode. A cortical reconstruction was generated from

the MRI using Freesurfer(93), and electrodes were localized using CT coregistration in Brainstorm(94). Seizure foci were identified in the bilateral mesial temporal lobes; however, the supratemporal plane and insulo-opercular regions were not affected. The sEEG signals were recorded with a Grapevine Nomad processor (Ripple LLC, Salt Lake City, UT) and theaccompanying Trellis recording software. The sEEG was recorded at a sampling rate of 1000 Hz and an online notch-filter was applied at 60/120/180 Hz to reduce electrical line interferences. The audio signal was synchronously recorded by the Grapevine system at a sampling rate of 30000 Hz. The auxiliary audio channel was used to mark the onset times of each stimulus in the sEEG recordings. Three electrode trajectories inserted along the superior temporal plane were considered for further analysis. These electrode trajectories were along the planum polare and planum temporale bilaterally.  The participant passively listened to the Mandarin tone and the click train stimuli. .

*Scalp Electroencephalography in Humans*

Scalp-recorded FFRs were recorded from the 20  human participants using EEG. The four Mandarin tones were presented using an alternating polarity at an intensity level of 72 dB SPL with interstimulus intervals jittered from 122 to 148 ms. EEG was recorded with a single AgCl electrode placed on the scalp that was referenced to the left mastoid, and the ground was placed on the opposite mastoid. Brainvision EEG system was used to record the EEG activity. A dedicated preamplifier (EP-preamp) connected to the actichamp amplifier with a gain setting of 50x was used. The EEG data was bandpass filtered from 80 to 1000 Hz and epoched from -25 to 250 ms (re:stimulus onset). Baseline correction was applied on each epoch, and the epochs exceeding an amplitude of +/-35 µV were excluded from further analysis.

## Intracranial and scalp electroencephalography in Rhesus Macaque

### *Subjects*

The EEG experiments were performed on three adult male macaque monkeys (macaca mulatta, animals B, W, J). Intracranial recordings in A1 were performed in a subset of two of these animals (laminar recording: animal W, single-tipped sharp electrode recordings: animal J). The treatment of the monkeys was in accordance with the guidelines set by the U.S. Department of Health and Human Services (National Institutes of Health) for the care and use of laboratory animals. All methods were approved by the Institutional Animal Care and Use Committee at the University of Pittsburgh. The animals were between 5 and 11 years old and weighed between 8-11 kg at the time of the experiments.

### *Cranial EEG recordings*

Details of the cranial EEG recordings have been reported previously (95,96). Briefly, EEG electrodes manufactured in-house from either medical grade stainless steel (monkeys J and W) or medical-grade titanium (monkey B) were implanted in 1mm deep, non-penetrating holes in the cranium. All electrodes were connected to a 32-channel Omnetics connector embedded in dental acrylic at the back of the skull. The different animals had 33 electrodes implanted that formed regularly-spaced grids covering roughly the same anatomy covered by the international 10-20 system (97). All the electrodes were referenced to an electrode placed at Oz.

### *Intracranial recordings in primary auditory cortex*

For the laminar recording, neural activity was recorded with a 24 channel laminar electrode (S-Probe from Plexon) positioned approximately perpendicular to the left superior temporal plane's orientation. The depth of the probe was adjusted iteratively until the prominent

sound-evoked supra-granular source was located slightly above the center of the probe. For the single-tipped sharp electrode recordings, neural activity was recorded with a chronically implanted 96 channel electrode array with individually movable electrodes (SC96 from Graymatter). The device was implanted over the right hemisphere in a way that allowed electrodes to approach the superior temporal plane approximately perpendicular. At the time of the experiments, 12 of the electrodes were positioned in or close enough to the superficial layers of the auditory cortex to pick up frequency-tuned local field potentials. Six of these electrodes also picked up frequency-tuned multi-unit activity, suggesting that they were located in layer III or below.

***Experimental Setup***

All experiments were performed in small (4’ wide by 4’ deep by 8’ high) sound-attenuating and electrically insulated recording booths (Eckel Noise Control Technology). Animals were positioned and head-fixed in custom-made primate chairs (Scientific Design). Neural signals were recorded with a 256-channel digital amplifier system (RHD2000, Intan) at a sampling rate of 30 kHz.

Experimental control was handled by a Windows PC running an in-house modified version of the Matlab software-package monkeylogic. Sound files were generated prior to the experiments and presented by a sub-routine of the Matlab package Psychtoolbox. The sound-files were presented using the right audio-channel of a high-definition stereo PCI sound card (M-192 from M-Audiophile) operating at a sampling rate of 96 kHz and 24 bit resolution. The analog audio-signal was then amplified by a 300 Watt amplifier (QSC GX3). The amplified electric signals were converted to sound waves using a single element 4 inch full-range driver (Tang Band W4-1879) located 8 inches in front of the animal. To determine sound onset with

high accuracy, a trigger signal was routed through the unused left audio channel of the sound card directly to one of the analog inputs of the recording system. The trigger pulse was stored in the same stereo sound-file and was presented using the same function call. Hence, any delay in the presentation of the tone also leads to an identical delay in the presentation of the trigger. Thus, sound onset could be determined at a level of accuracy that was limited only by the sampling frequency of the recording device (30kHz: corresponding to 33 μsec).

## Cranial electroencephalography and intracranial recordings in guinea pigs

### *Subjects*

The EEG recordings were performed on two female and a male adult (~ 8 months old), wild-type, pigmented guinea pigs (GPs; Cavia porcellus; Elm Hill Labs, Chelmsford, MA), weighing ~600 – 800 g. Intracranial recording from A1 was performed in the male GP. All experimental procedures were conducted according to NIH Guidelines for the care and use of laboratory animals, and were approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Pittsburgh.

Prior to commencing recordings, a custom headpost for head fixation, skull screws that served as EEG recording electrodes or reference electrodes for intracranial recordings, and recording chambers for intracranial recordings were surgically implanted onto the skull using dental acrylic (C & B Metabond, Parkell Inc.) following aseptic techniques under isoflurane anesthesia. Analgesics were provided for three days after surgery, and animals were allowed to recover for ~10 days. Following recovery, animals were gradually adapted to the recording setup and head fixation for increasing durations of time.

### *Experimental Setup*

All recordings were performed in a sound-attenuated booth (IAC) whose walls were covered with anechoic foam (Pinta Acoustics). Animals were head-fixed in a custom acrylic enclosure affixed to a vibration isolation tabletop. Stimuli were presented using Matlab (Mathworks, Inc). Digital stimulus files sampled at 100 kHz were converted to an analog audio signal (National Instruments, USA), attenuated (Tucker-Davis Technologies, USA), power-amplified (Tucker-Davis Technologies, USA) and delivered through a calibrated speaker (4" full-range driver, TangBand, Taiwan) located ~0.9 m in front of the animal. Stimuli were presented at ~75 dB SPL.

### *Cranial EEG recordings*

FFRs were acquired from unanesthetized, head-fixed, passively-listening GPs using a vertical electrode montage. Scalp-recorded activity was collected via a stainless-steel skull screw (Fine Science Tools, USA). Reference and ground conductive adhesive hydrogel electrodes (Foam electrodes, Kendall™ or Covidien Medi-trace®, USA) were placed on the mastoids. Signals were acquired using a multichannel neural processor (Ripple Inc., USA).

### *Intracranial recordings in primary auditory cortex*

Intracortical recordings were performed in unanesthetized, head fixed, passively-listening GPs. Small craniotomies (1 – 2 mm. Diameter) were performed within the implanted recording chambers, over the expected anatomical location of A1. Neural activity was recorded using high-density 64-channel multi-site electrode (Cambridge Neurotech), inserted roughly perpendicular to the cortical surface. The tip of the probe was slowly inserted to a depth of ~2 mm. After a

brief waiting period to allow the tissue to settle, signals were acquired using a multichannel neural processor.

***FFR Preprocessing and Analysis for sEEG***

A linear regression based method was used to remove the harmonics of the power line interference from the data, using the cleanline plugin in EEGLAB(98). The raw sEEG was high pass filtered using a third-order zero-phase shift butterworth filter. No low pass filter was used, as the sampling rate was 1000 Hz resulting in an effective low pass frequency of 500 Hz. Time-locked sEEG epochs were extracted for all vowels in both polarities. The epochs that exceeded amplitudes of 75 µV were rejected. The FFRs in both the polarities for each vowel were averaged to obtain a total of four FFR waveforms (one for each vowel). Four FFR waveforms were obtained for each electrode.

The phase locking value (PLV) was estimated to assess the frequencies at which the cortical units phase-lock. The single -rial FFRs were decomposed into a spectrogram representation using time-frequency analysis timefreq.m in EEGLAB. Specifically, 130 wavelets between 70 and 200 Hz with equal widths were used for the time-frequency decomposition. The complex valued time-frequency vectors were divided by their magnitude to obtain unit vectors at every frequency and time point. These unit vectors in the time-frequency domain were averaged across trials to obtain resultant vectors. The absolute magnitude of the resultant vectors at each time and frequency point was used to obtain the PLV spectrogram (Figure 1). PLVs ranged from 0 to 1, with 0 indicating no phase-locking and 1 indicating perfect phase-locking across trials. The PLVs provide information about the extent of phase-locking at different frequencies and latencies without the confounds of differences in absolute FFR magnitude.

**FFRs to Mandarin tones in macaques and guinea pigs**

The raw data were high pass filtered using a second-order zero-phase shift FIR filter. Time-locked epochs were extracted for all vowels in both polarities. The epochs that exceeded amplitudes of 250 µV were rejected. The FFRs in both the polarities for each tone were averaged to obtain a total of four FFR waveforms (one for each tone). Current source densities were calculated as the second spatial derivative following spatial smoothing.

Custom Matlab and R routines were used to filter and average signals appropriately to obtain local field potentials (LFPs) and multi-unit activity (MUA) from the macaque and GP recordings. The current source density (CSD) in the macaque was computed from the LFP signals derived from the electrodes with 150 µm spacing, spatially smoothing the LFP signal using a Gaussian filter (SD = 250 µm), and obtaining the second spatial method using the finite difference approximation. In the GP, the electrodes spaced 60 µm apart were subsampled to approximate the macaque recordings. The CSD was computed from the LFP signals derived from alternate electrode contacts (60 µm spacing), spatially smoothing the LFP signal using a Gaussian filter (SD = 125 µm), and obtaining the second spatial method using the finite difference approximation.

**Representational similarity analysis of FFRs to Mandarin tones**

A hidden Markov model (HMM) classifier was trained to decode the stimulus pitch patterns based on the FFR pitch tracks(52). A detailed description of the HMM-based decoding approach can be found in a previously published methods paper(52). The averaging size of the HMM was dynamically adjusted to obtain equivalent classification accuracies across the different levels (scalp, cortex) and species. The averaging sizes used were 150 trials for human scalp FFR, 24 trials for human sEEG FFR, 150 trials for macaque scalp FFR, 4 trials for

macaque PAC FFR, 6 trials for the Guinea pig scalp FFRs, and 16 trials for the GP PAC FFRs. The confusion matrices of decoding patterns (proportion correct) were extracted and used for further representational similarity analysis. Multi-dimensional scaling analysis was performed on the confusion matrices (diagonals removed) to assess if the decoding patterns across levels and species converge in the same space. Procrustes analysis was performed to rotate and transform all the MDS representation to the same scale to facilitate comparison. A similarity matrix was derived from the confusion matrices (diagonals removed) across-levels and-species using Pearson's product-moment correlations.

**Comparison of FFRs recorded at the scalp and the auditory cortex**

*Spectro-temporal Measures*

In the animal models we had the opportunity to recorded FFRs from the scalp and the cortex in the same animal. We compared the FFR power spectra at the scalp and cortex to analyze similarities and differences in the spectral composition across the two units of analyses. Welch's power spectral density (PSD) estimates of the FFRs were obtained with a 1024 point hamming window with 50% overlap, to obtain a smoothed spectral estimate of the FFRs. The PSDs of all the four mandarin tones were averaged to obtain an average spectral composition of the FFRs. The PSD of the scalp and cortical FFRs were normalized by setting their maximum magnitude to 0 dB. This facilitated the comparison of the spectral decay relative differences in encoding of the F0 and the higher harmonics at the scalp and the cortex. Due to the high amplitude of the FFR at the F0, the normalization process essentially normalized the F0 magnitude, which facilitated the inspection of decay in magnitude of the high frequency components in the FFRs with reference to the F0 magnitude.

Due to the use of the varying F0 trajectories in the stimulus, it is challenging to infer the FFR temporal properties from a singular estimate of cross-correlation latency based on the raw FFR waveforms. This is especially challenging when the underlying sources of FFRs may change depending on the frequency. Thus we, chose to compare the latencies between the scalp and cortical FFRs in the time-frequency domain. We decomposed the FFRs into a time-frequency representation using the continuous wavelet transformation using the Morse wavelet implementation in the Matlab wavelet analysis toolbox. The real-valued wavelet decomposed waveforms at every frequency were then cross-correlated between the scalp and cortical FFRs. The cross-correlation lags with the highest absolute correlation coefficient were estimated. This analysis provided a latency estimate (cross-correlation lag) and magnitude of similarity (Pearson's *r*) for every frequency in the FFR. The latencies measured here are with the scalp FFRs as reference. Thus, positive latencies indicate that the cortical FFRs lag scalp FFRs in latency. These metrics were obtained both in the Macaque and Guinea Pig FFRs. In the Macaque, scalp recording was obtained simultaneously with FFRs from the cortical surface in the same session. One scalp and one cortical electrode with the best signal-noise ratio were chosen for this analysis. In the Guinea pig, FFRs were obtained in the two different sessions from the scalp and the cortex. The FFRs from the cortex were derived from the electrode with the maximum signal-to-noise ratio in laminar FFR-LFP recordings (Electrode 10 in Layer 3).

A cross-spectral power density estimate was also obtained to assess the similarity in power between the scalp and cortical FFRs irrespective of the latency difference. This estimate is useful in getting an objective metric of similarity in spectral properties of the FFRs at the cortex and the scalp, irrespective of the differences in temporal properties. The Cross-spectral densities were obtained using Welch's periodogram method with 50% overlapping 1024-point hamming

windows. The absolute power of the cross-spectral densities was averaged across the FFRs for the four Mandarin tones and overlaid on the plot of frequency-wise latency comparison to obtain a unified inference of the spectro-temporal similarities in the FFRs at the scalp and cortex.

*Independent Component Analysis*

The above measures show differences and similarities in the scalp and cortically recorded FFRs. Multiple volume-conducted fields in the cortex and sub-cortex that lead to constructive and destructive interferences drive these differences and similarities, and the above measures may not be sensitive to differentiate these fields. Thus, we further used a  blind source separation approaches using independent component analysis that can separate spectro-temporally overlapping components arising from different neural sources(99). ICA extracts mutually independent components disentangling the superposed electrical fields across the different electrical dipoles across the auditory cortical subcortical regions and their projection to the scalp and the cortical electrodes. The electrodes placed at the cortex though record activity from the cortical ensembles in their proximity, they also pick up electrical fields from the subcortical regions and the surrounding auditory cortical fields. With ICA, we can decompose the neural sources underlying the electrical activity across the cortical and scalp electrodes and trace the differential projection of neural sources to the individual electrodes.

ICA was applied to the single-trial data from 33 scalp electrodes and 96 sharp electrodes from the cortical surface (spanning the surface of superior temporal plane and motor cortices) in the Macaque, and 2 scalp channels (Cz: the intersection of the midline sagittal and coronal planes, T4: on the right temporal bone above the auditory cortex) and 25 laminar depth electrodes in the auditory cortex  of the Guinea Pig. ICA was done separately for both species. The averaged FFRs corresponding to all four Mandarin tones were input into the ICA

decomposition. The Infomax algorithm (runica.m) in EEGLAB toolbox(98) was used for ICA decomposition of the neural data. A principal component analysis was performed to reduce the dimensionality of the signal that explained 96% (Macaque) and 99% (Guinea Pig) of the original variance in the data. The percentage variance accounted for by each independent component (IC) was estimated (eeg_pvaf.m). The ICs that each explained greater than 10% of the variance in the FFRs were retained for further analysis. The spatial weights of the ICAs were derived as the pseudoinverse of the product of the ICA weights and the ICA sphering matrix. These spatial weights were used to assess the spatial layout of the volume-conducted propagation of the different ICAs. The PVAFs of the ICAs at each scalp and cortical electrode were estimated to obtain the contribution of the ICAs to each of the electrodes. The emphasis of this analysis was to assess the percent contribution of the cortical ICAs to the scalp-recorded FFRs. The latencies of the ICAs were assessed by cross-correlation with the stimulus waveform. These latencies were also used to infer the potential generators of the ICAs, with earlier latencies corresponding to more subcortical sources. The power coherence was estimated between each of these ICAs and the stimulus waveform using cross-power spectral density (cpsd.m). This analysis aided in inferring the differential pattern in the decline of power coherence across the cortical ICAs and subcortical ICAs.  The power coherence metric is not a measure of phase-locking but just the power coherence between the stimulus and the ICAs.

## ACKNOWLEDGEMENTS

This investigation was supported by National Institute of Health funding to R01-DC013315 to B.C., S.S., T.A, and T.T; RF1-MH114223 to T.T., and R01-DC017141 to S.S. This work was also supported in part by the resources provided through the University of Pittsburgh Center for Research Computing. Preliminary findings from this investigation were presented in the symposium of Advances and Perspectives in Auditory Neuroscience 2020.

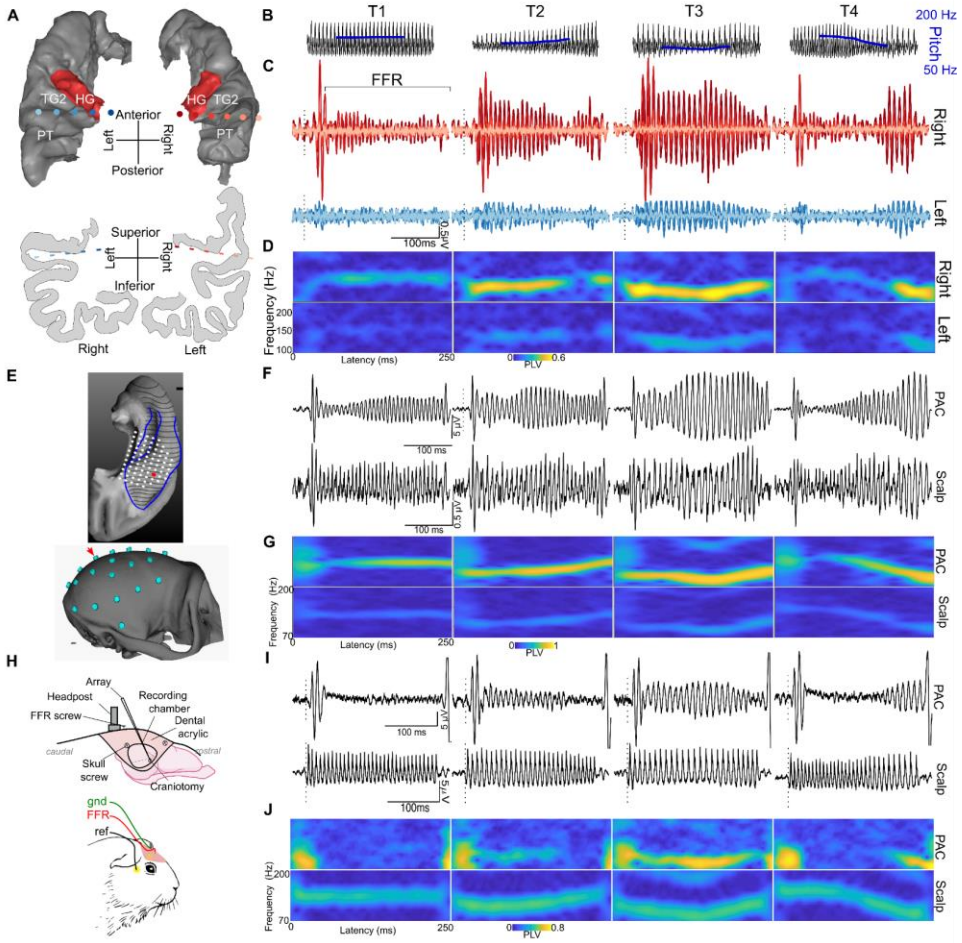



**Figure 1. Cross-species and cross-level approach to characterizing the functional properties of cortical FFRs**. **A-D** human data, **E-G** Macaque data, **H-J** Guinea pig data. **A** Superior view *(Top)* of bilateral temporal lobes showing the location of the sEEG electrodes (projected to the surface for visualization) in the human participant. Coronal view *(Bottom)* of the brain showing the location of the sEEG electrodes. The slice is adjusted perpendicular to the axis of the electrode shaft. **B** Waveforms of the Mandarin tones /yi/; T1 (high-level fundamental frequency (F0)), T2 (low-rising F0), T3 (low-dipping F0), and T4 (high-falling F0). **C** FFR waveforms from sEEG electrodes in both right (top panel – right) and left (bottom panel - blue) temporal lobes in the human participant. The waveforms with the darkest color are from the electrodes located medially along the temporal lobe and closer to the Heschl's gyrus (HG), while the waveforms with the lightest colors are from electrodes located laterally along the temporal lobe closer to the planum temporale (PT) and superior temporal gyrus (STG). **D**: FFR phase-locking consistency across trials at all time points and frequencies with a spectrogram showing the phase-locking value (PLV). The PLV spectrograms shown are only for the electrodes with the highest FFR amplitude in each hemisphere (located on the HG). **E** *Top* Locations of the semi-chronic sharp electrodes (white dots) used to record FFRs in the primary auditory cortex (PAC) of the Macaque. FFRs from a representative example electrode highlighted in red. *Bottom* Layout of the EEG Electrode grid. FFR responses analyzed here were from the electrode marked with the red arrow. **F** Waveforms show simultaneously recorded FFRs from Macaque primary auditory cortex (Top - PAC) and skull (Bottom - ~ Cz). **G** PLV spectrograms of FFRs in the Macaque at

both PAC and scalp. **H** intracranial and extracranial electrode setup in the Guinea Pig. **I** Waveforms show FFRs to the four Mandarin tones recorded from the best laminar depth electrode placed in the PAC (*top*) and from a surface scalp electrode (*bottom*) of the Guinea Pig. **J** PLV spectrograms of FFRs in guinea pigs.

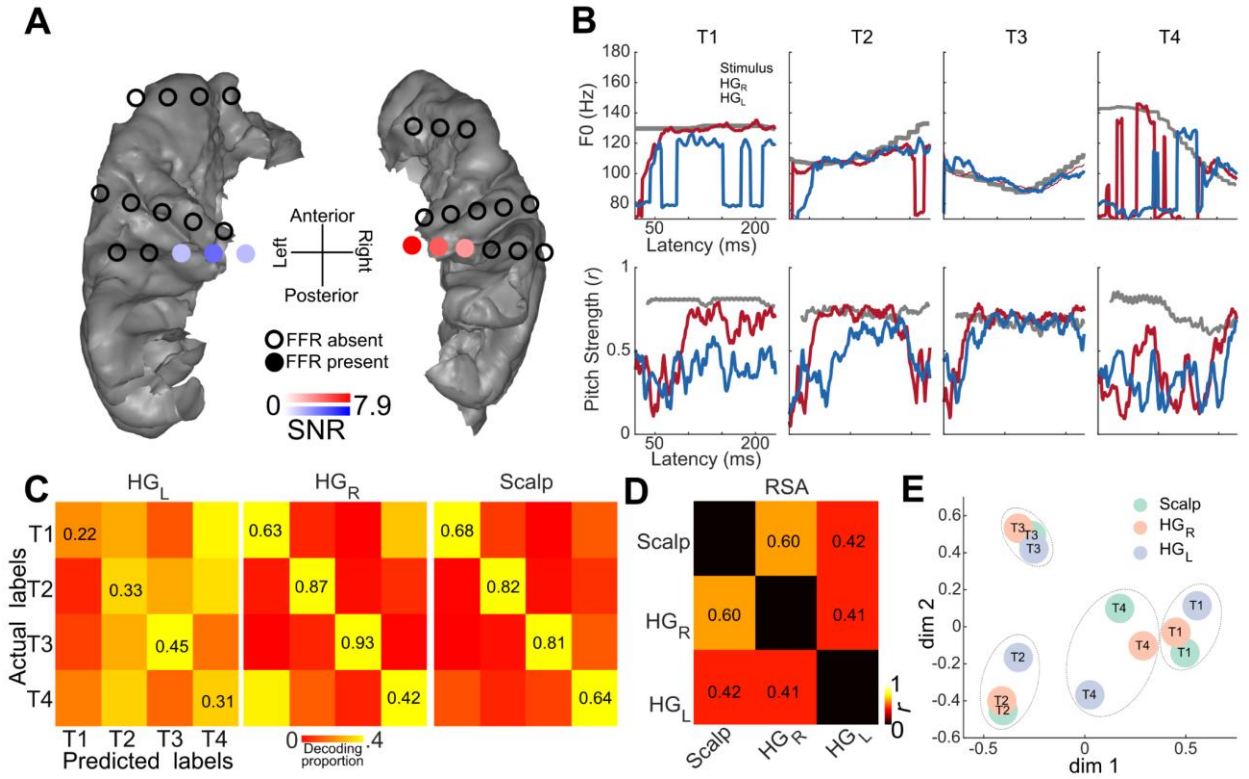



**Figure 2. Hemispheric differences in the frequency following responses (FFRs). A.** Electrodes implanted in the superior temporal plane are shown with filled circles marking significant FFR root mean square amplitude above the pre-stimulus baseline ($p<0.05$ permutation-based bootstrapped t-tests) and unfilled circles mark electrodes where the FFRs were not significantly above the baseline. **B.** Pitch tracking measures (sliding window autocorrelation analysis) of the FFRs from the best electrodes in the two hemispheres showing the hemispheric differences. Top row shows the derived F0 track in Hz (best lag in each sliding window) in the right ($HG_R$) and left ($HG_L$) heschl's gyri. Bottom row shows the pitch strength (maximum correlation coefficient in each sliding window). **C.** Hidden markov model based decoding accuracies of the Mandarin tone from the FFRs recorded in the right and left Heschl's gyri from the sEEG recording, scalp EEG recordings from 20 different participants. **D.** Representational similarity (Pearson's correlation) between the FFR confusion matrices (diagonals removed) in C.The correlation $HG_R$ and scalp FFR was significant ($p=0.039$) while the other correlations were not significant. **E.** Multi-dimensional scaling of the confusion matrices in C, visualizing the confusion space for decoding the Mandarin tone.

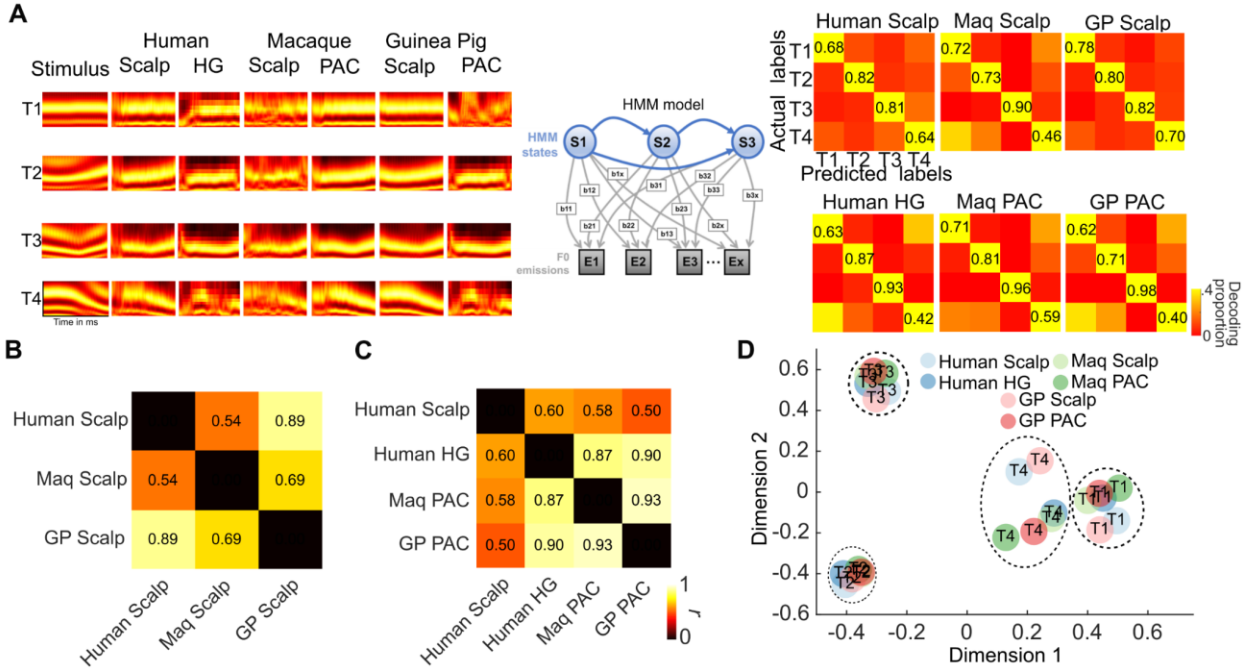



**Figure 3. Cross-species and cross-level representational similarity analyses of FFRs. A** Hidden Markov Model (HMM) was used to assess the extent to which pitch patterns (T1-High flat, T2-Low rising, T3-Low dipping, and T4-High falling) could be decoded from the FFRs. HMM decoding accuracies were estimated for each species (Human, Macaque, and Guinea pig) and each level (scalp and intracortical). HMM decoding accuracies were significantly above chance for all species, levels. Confusion matrices *(right)* show the accuracies for decoding (along the diagonal) the pitch patterns from FFRs and error patterns. The averaging size used for decoding was adjusted to obtain comparable classification accuracies at the scalp and the cortex. The confusion matrices in A (principal diagonal removed) were assessed for correlations (Pearson's *r*) between species and between levels to estimate the cross-species and cross-level representational similarity. **B.** shows the correlation between the confusion matrices for scalp FFR across species. FFRs recorded at the scalp share strong similarities ($p<0.05$). However, the correlation between human and Macaque scalp-recorded FFR did not reach statistical significance (p=0.06). **C.** shows the correlation between the confusion matrices of intracortical FFRs across species. Intracortical FFRs showed very strong correlations ($p<0.05$). Also shown is the correlation of intracortical FFRs and the scalp-recorded FFRs in humans. across species. The human scalp FFRs and GP intracortical FFRs did not show a significant correlation (p = 0.09). . **D.** Multi-dimensional scaling (MDS) analysis of the pitch patterns based on the confusion matrices in each species and level.

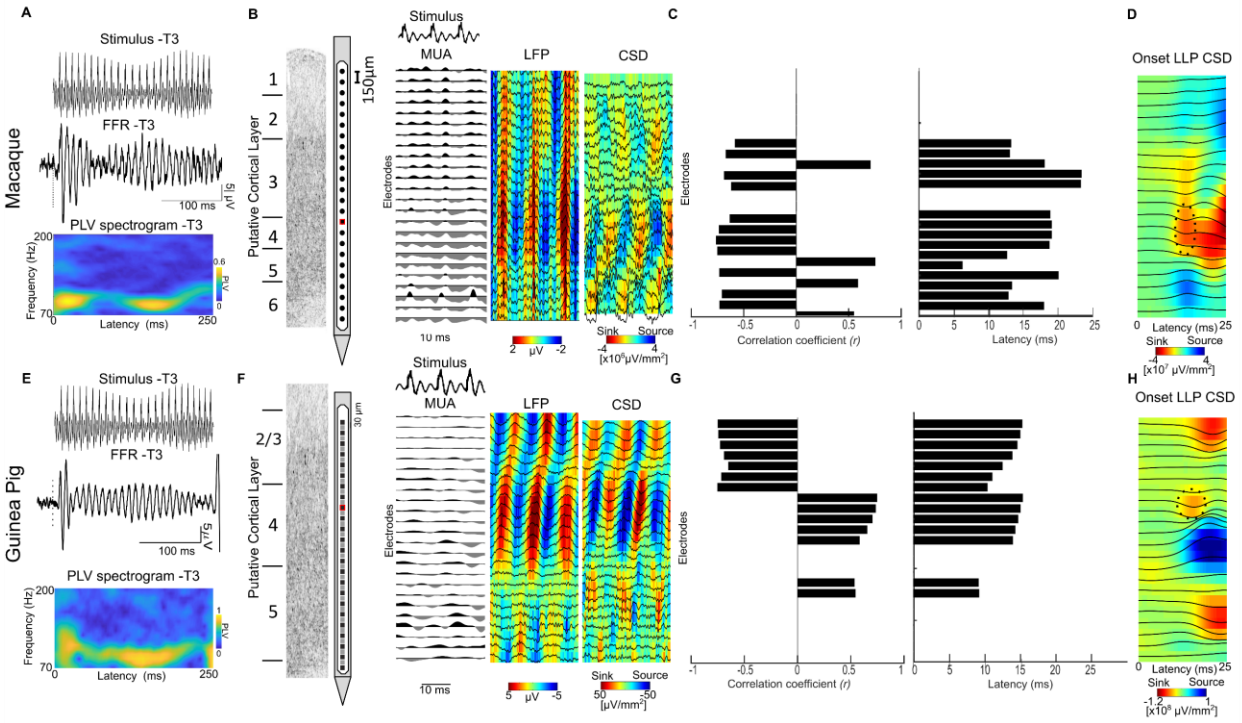



**Figure 4. Cortical source for FFRs to speech syllables in the Macaque and Guinea Pig models**: FFRs recorded in the Rhesus Macaque and Guinea Pig models using depth electrodes. **A** and **E** show sample stimulus (T3- low frequency dipping contour) and FFRs with the best signal to noise ratios in the putative layer IV in the laminar probe. Waveform shown is from the electrode with maximum amplitude. Also shown is the phase-locking value spectrogram for the FFR (electrode shown in red in B and F). **B** and **F** show FFRs to stimulus T3, recorded using multi-channel laminar-depth electrodes in the Macaque and the Guinea pig, respectively. The FFRs are shown only for a short segment of the stimulus to clearly visualize the patterns. The multi-unit activity (MUA) and local field potentials (LFP) show strong phase-locked activity to each periodicity cycle in the stimulus. The source-sink configuration in the reference-free current source densities (CSD) show existence of an FFR generator in putative Layers III and IV of primary auditory cortex. **C** and **G** shows the stimulus to responses correlation coefficients and latencies (lags) between the stimulus and the FFR CSDs in each layer. The shift in polarity of the CSDs is manifested as a shift in the sign of correlation coefficient and shows sources of FFRs in two cortical layers. **D** and **G** show the current source densities (CSD) of the long latency potentials (LLPs) to the onset of the Mandarin tones. The first sink (dotted ellipse) in the CSD marks the putative cortical layer IV.

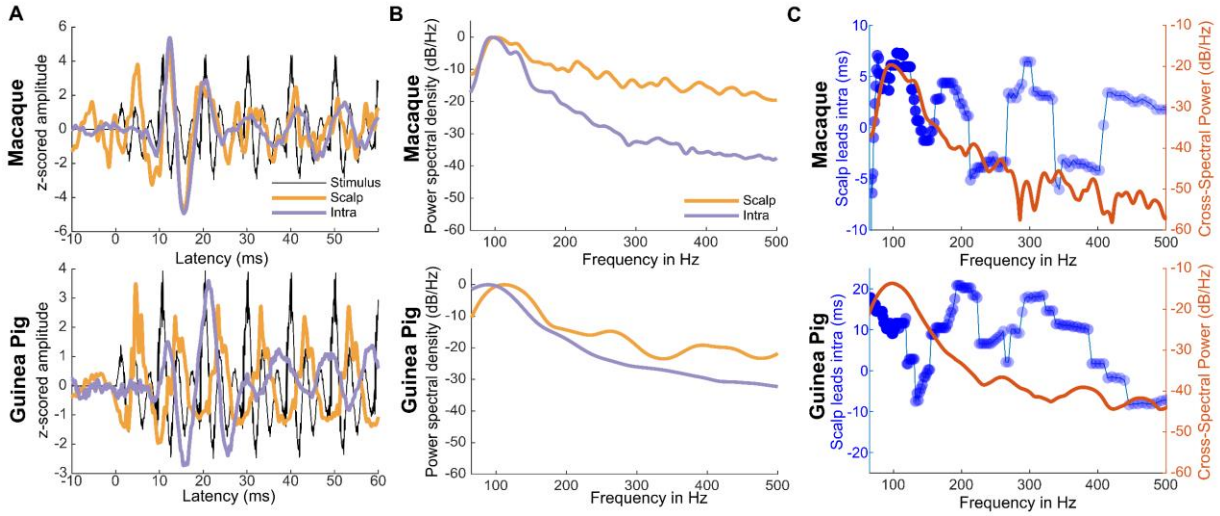



**Figure 5. Scalp and intracortical FFRs in the Macaque and the Guinea Pig. A.** FFR waveforms to Mandarin tone 3 (low dipping F0) at scalp and the cortex in the Macaque and the Guinea Pig. **B.** Normalized power spectral density of the scalp and intracortical FFRs in the Macaque and Guinea Pig, showing the different the low-frequency dominance of the intracortical FFRs. **C.** Difference between scalp-recorded and intracortical FFRs. Blue tracings show the latency between scalp and intra-cortically recorded FFRs. Circles with darker colors (colors normalized to maximum correlation across frequencies)) indicate a higher correlation between the scalp and intra-cortically recorded FFRs. Intracortical FFRs showed longer latencies than scalp-recorded FFRs at frequencies below 120 Hz. Red tracings show the cross-spectral density between scalp and intra-cortically recorded FFRs showing shared power at low frequencies.

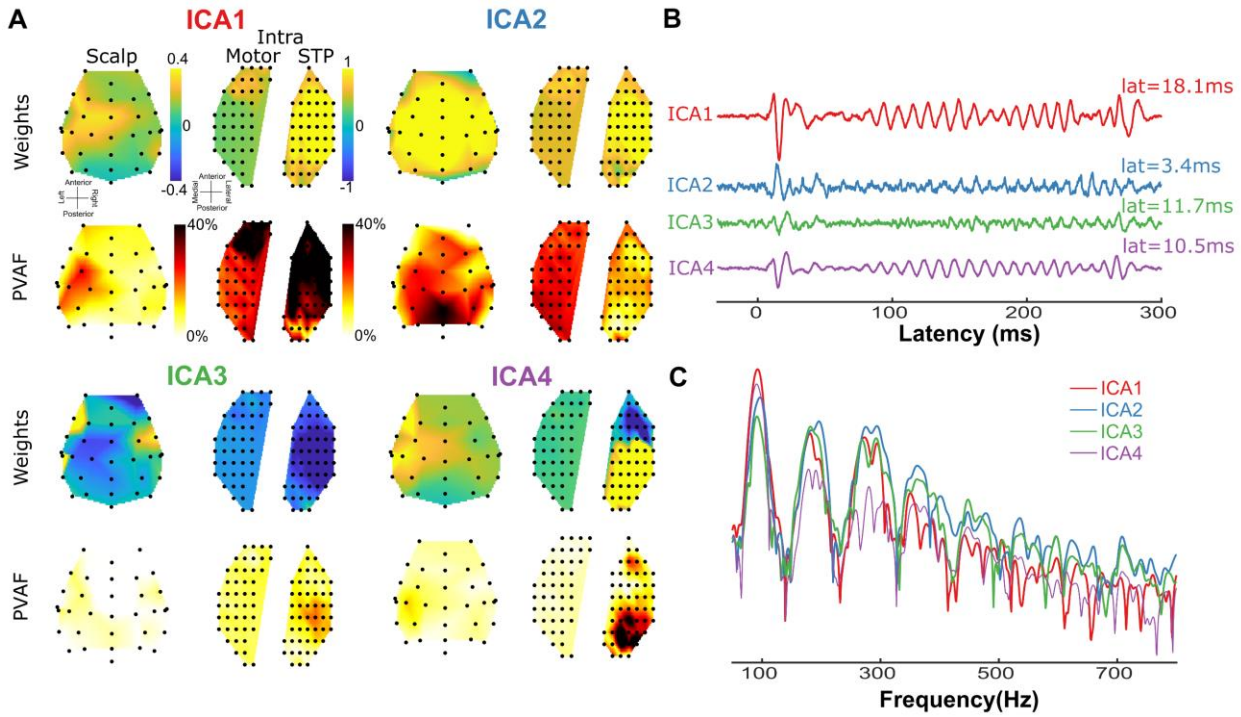



**Figure 6. Evaluating the contribution of intracortical FFRs to scalp FFRs using independent component analysis (ICA) in the Macaque. A.** Spatial loadings (ICA weights) of the top 4 independent components onto the scalp and intracortical (Intra electrodes located in the motor regions and in the superior temporal plane) electrodes and the percentage variance accounted for (PVAF) by the ICAs at each electrode. **B.** Time course of the activations of the top six ICs (amplitude in arbitrary units) for stimulus T3 (low dipping F0). **C.** Power spectral coherence of the top six ICAs contributing to the FFRs(amplitude in arbitrary units).

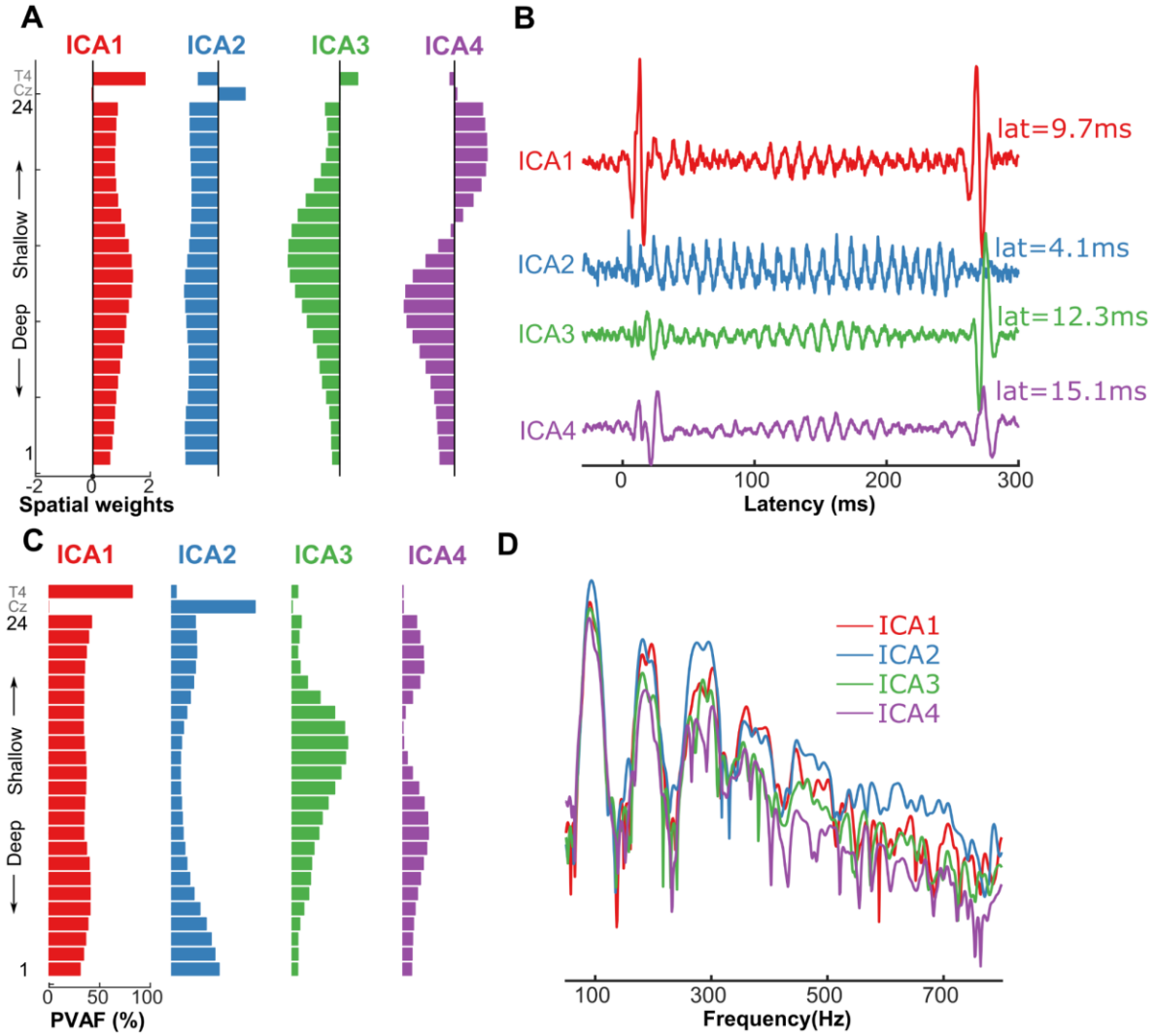



**Figure 7. Evaluating the contribution of intracortical FFRs to scalp FFRs using independent component analysis (ICA) in the GP. A.** Spatial loadings of the top 4 independent components onto the intracortical laminar electrodes (1-24) and scalp electrodes (Cz and T4). **B.** Time course of the activations of the top four ICs (amplitude in arbitrary units). **C.** Spatial loadings of the top six independent components onto the intracortical laminar electrodes (1-24) and scalp electrodes (Cz and T4). **D.** Power spectral coherence of the top four ICAs contributing to the FFRs (amplitude in arbitrary units).

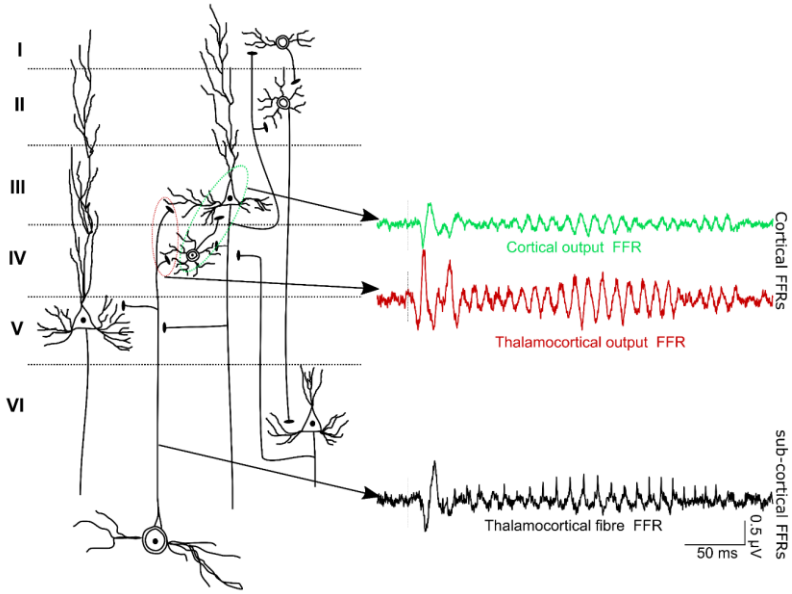
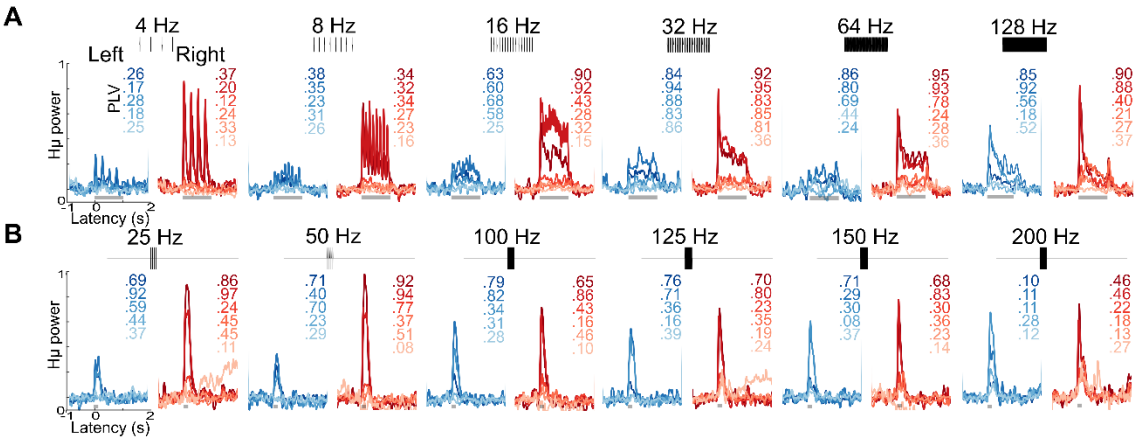



**Figure 8.** Schematic representation of the cortical FFRs and the subcortical FFRs. Any phase locked currents in the cortical layers are considered the cortical FFRs. These also consist the thalamocortical outputs that synapse with the layers III and IV,and generate post-synaptic potentials that are phase-locked to the stimulus F0.




**Supplementary Figure 1. Frequency following responses to click train stimuli**. **A** and **B** show the event-related high gamma power responses (Hμ) to the two different sequences of click trains presented at different repetition rates. The black tracings illustrate the stimulus waveform at each of the click repetition rates. The click trains in **A** were 1000 ms long, while the click trains in **B** were 160 ms long. A gray shaded bar below each Hμ waveform shows marks the duration of the stimulus. The numbers shown as insets represent the phase-locking value (PLV) at click repetition rate obtained from time-frequency analysis of single-trial local-field potentials. *Note: The waveforms/numbers with the darkest color are from the electrodes located medially along the temporal lobe and closer to the Heschl's gyrus (HG), while the waveforms with the lightest colors are from located laterally along the temporal lobe and closer to the planum temporale (PT) and superior temporal gyrus (STG).*

## References


1. Plack CJ, Barker D, Hall DA. Pitch coding and pitch processing in the human brain. Hearing Research. 2014 Jan;307:53–64.

2. Langner G. Periodicity coding in the auditory system. Hearing Research. 1992 Jul 1;60(2):115–42.

3. Walker KMM, Bizley JK, King AJ, Schnupp JWH. Cortical encoding of pitch: Recent results and open questions. Hear Res. 2011 Jan;271(1–2):74–87.

4. Ayala YA, Lehmann A, Merchant H. Monkeys share the neurophysiological basis for encoding sound periodicities captured by the frequency-following response with humans. Scientific Reports. 2017 Nov 30;7(1):1–11.

5. Smith JC, Marsh JT, Brown WS. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol. 1975 Nov;39(5):465–72.

6. Marsh JT, Brown WS, Smith JC. Far-field recorded frequency-following responses: correlates of low pitch auditory perception in humans. Electroencephalogr Clin Neurophysiol. 1975 Feb;38(2):113–9.

7. Greenberg S, Marsh JT, Brown WS, Smith JC. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hearing research. 1987;25(2):91–114.

8. Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. Ear Hear. 2010 Jun;31(3):302–24.

9. Krizman J, Kraus N. Analyzing the FFR: A tutorial for decoding the richness of auditory function. Hearing Research. 2019 Oct;382:107779.

10. Gnanateja GN, Ranjan R, Sandeep M. Physiological bases of the encoding of speech evoked frequency following responses. Journal of All India Institute of Speech and Hearing. 2012;31:215–9.

11. Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology. 2010 Mar;47(2):236–46.

12. Coffey EBJ, Nicol T, White-Schwoch T, Chandrasekaran B, Krizman J, Skoe E, et al. Evolving perspectives on the sources of the frequency-following response. Nat Commun. 2019 Nov 6;10(1):1–10.

13. White-Schwoch T, Carr KW, Thompson EC, Anderson S, Nicol T, Bradlow AR, et al. Auditory Processing in Noise: A Preschool Biomarker for Literacy. PLOS Biology. 2015 Jul 14;13(7):e1002196.

14. Gockel HE, Carlyon RP, Mehta A, Plack CJ. The Frequency Following Response (FFR) May Reflect Pitch-Bearing Information But is Not a Direct Representation of Pitch. Journal of the Association for Research in Otolaryngology. 2011 Dec;12(6):767–82.

15. Gnanateja GN, Maruthy S. Dichotic phase effects on frequency following responses reveal phase variant and invariant harmonic distortion products. Hearing Research. 2019 Sep;380:84–99.

16. Gerken GM, Moushegian G, Stillman RD, Rupert AL. Human frequency-following responses to monaural and binaural stimuli. Electroencephalogr Clin Neurophysiol. 1975 Apr;38(4):379–86.

17. Worden FG, Marsh JT. Frequency-following (microphonic-like) neural responses evoked by sound. Electroencephalogr Clin Neurophysiol. 1968 Jul;25(1):42–52.

18. Gardi J, Merzenich M, McKean C. Origins of the Scalp-Recorded Frequency-Following Response in the Cat. Audiology. 1979 Jan 1;18(5):353–80.

19. Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N. Context-Dependent Encoding in the Human Auditory Brainstem Relates to Hearing Speech in Noise: Implications for Developmental Dyslexia. Neuron. 2009 Nov;64(3):311–9.

20. Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. PNAS. 2009 Aug 4;106(31):13022–7.

21. Russo N, Nicol T, Trommer B, Zecker S, Kraus N. Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Developmental Science. 2009 Jul;12(4):557–67.

22. Abrams DA, Kraus N. Biomap: A neurodiognostic tool for auditory processing disorders. The ASHA Leader [Internet]. 2005 [cited 2016 Jan 22]; Available from: http://www.asha.org/Publications/leader/2005/051018/f051018c/

23. Hornickel J, Kraus N. Unstable Representation of Sound: A Biological Marker of Dyslexia. Journal of Neuroscience. 2013 Feb 20;33(8):3500–4.

24. Billiet CR, Bellis JT. The Relationship Between Brainstem Temporal Processing and Performance on Tests of Central Auditory Function in Children With Reading Disorders. Journal of Speech, Language, and Hearing Research. 2011 Feb 1;54(1):228–42.

25. Breier JI, Fletcher JM, Foorman BR, Klaas P, Gray LC. Auditory temporal processing in children with specific reading disability with and without attention deficit/hyperactivity disorder. Journal of Speech, Language, and Hearing Research. 2003;46(1):31–42.

26. Kraus N, Thompson EC, Krizman J, Cook K, White-Schwoch T, LaBella CR. Auditory biological marker of concussion in children. Sci Rep [Internet]. 2016 Dec 22 [cited 2020 Jun 26];6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5178332/

27. Vander Werff KR, Rieger B. Brainstem Evoked Potential Indices of Subcortical Auditory Processing Following Mild Traumatic Brain Injury. Ear Hear. 2017;38(4):e200–14.

28. Kraus N, Lindley T, Colegrove D, Krizman J, Otto-Meyer S, Thompson EC, et al. The neural legacy of a single concussion. Neuroscience Letters. 2017 Apr 12;646:21–3.

29. Alain C, Zendel BR, Hutka S, Bidelman GM. Turning down the noise: The benefit of musical training on the aging auditory brain. Hearing Research. 2014 Feb;308:162–73.

30. Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging Affects Neural Precision of Speech Encoding. Journal of Neuroscience. 2012 Oct 10;32(41):14156–64.

31. Maruthy S, Kumar UA, Gnanateja GN. Functional Interplay Between the Putative Measures of Rostral and Caudal Efferent Regulation of Speech Perception in Noise. Journal of the Association for Research in Otolaryngology. 2017 Aug;18(4):635–48.

32. Bidelman GM, Villafuerte JW, Moreno S, Alain C. Age-related changes in the subcortical-cortical encoding and categorical perception of speech. Neurobiol Aging. 2014 Nov;35(11):2526–40.

33. Kraus N, Anderson S. The Effects of Aging on Auditory Processing. The Hearing journal. 2013;66(1):36.

34. Presacco A, Simon JZ, Anderson S. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. Journal of Neurophysiology. 2016 Aug 17;116(5):2346–55.

35. Galbraith GC, Threadgill MR, Hemsley J, Salour K, Songdej N, Ton J, et al. Putative measure of peripheral and brainstem frequency-following in humans. Neuroscience letters. 2000;292(2):123–7.

36. Marsh JT, Brown WS, Smith JC. Differential brainstem pathways for the conduction of auditory frequency-following responses. Electroencephalogr Clin Neurophysiol. 1974 Apr;36(4):415–24.

37. Yamada O, Marsh RR, Potsic WP. Generators of the Frequency-following Response in the Guinea Pig. Otolaryngol Head Neck Surg. 1980 Sep 1;88(5):613–8.

38. Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ. Cortical contributions to the auditory frequency-following response revealed by MEG. Nat Commun. 2016 Mar 24;7:11070.

39. Coffey EBJ, Arseneau-Bruneau I, Zhang X, Baillet S, Zatorre RJ. Oscillatory entrainment of the Frequency Following Response in auditory cortical and subcortical structures. J Neurosci [Internet]. 2021 Mar 17 [cited 2021 Apr 13]; Available from: https://www.jneurosci.org/content/early/2021/03/16/JNEUROSCI.2313-20.2021

40. Hartmann T, Weisz N. Auditory cortical generators of the Frequency Following Response are modulated by intermodal attention. NeuroImage. 2019 Dec 1;203:116185.

41. Coffey EBJ, Chepesiuk AMP, Herholz SC, Baillet S, Zatorre RJ. Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. Front Neurosci [Internet]. 2017 Aug 25 [cited 2018 Mar 30];11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5575455/

42. Greenberg RP, Stablein DM, Becker DP. Noninvasive localization of brain-stem lesions in the cat with multimodality evoked potentials: Correlation with human head-injury data. Journal of Neurosurgery. 1981 Jun 1;54(6):740–50.

43. Bidelman GM. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. NeuroImage. 2018 Jul 15;175:56–69.

44. Galbraith GC, Bagasan B, Sulahian J. Brainstem frequency-following response recorded from one vertical and three horizontal electrode derivations. Percept Mot Skills. 2001 Feb 1;92(1):99–106.

45. King A, Hopkins K, Plack CJ. Differential Group Delay of the Frequency Following Response Measured Vertically and Horizontally. JARO. 2016 Feb 26;17(2):133–43.

46. Bidelman GM. Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. Hearing Research. 2015 May;323:68–80.

47. Behroozmand R, Oya H, Nourski KV, Kawasaki H, Larson CR, Brugge JF, et al. Neural Correlates of Vocal Production and Motor Control in Human Heschl's Gyrus. J Neurosci. 2016 Feb 17;36(7):2302–15.

48. Guo N, Si X, Zhang Y, Ding Y, Zhou W, Zhang D, et al. Speech frequency-following response in human auditory cortex is more than a simple tracking. NeuroImage. 2020 Nov 11;117545.

49. Krishnan A, Gandour JT, Bidelman GM. Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. NeuroReport. 2012 May;23(8):498–502.

50. Reetzke R, Xie Z, Llanos F, Chandrasekaran B. Tracing the Trajectory of Sensory Plasticity across Different Stages of Speech Learning in Adulthood. Current Biology. 2018 May;28(9):1419-1427.e4.

51. Wong PCM, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci. 2007 Apr;10(4):420–2.

52. Llanos F, Xie Z, Chandrasekaran B. Hidden Markov modeling of frequency-following responses to Mandarin lexical tones. Journal of Neuroscience Methods. 2017 Nov 1;291:101–12.

53. Lau JCY, Wong PCM, Chandrasekaran B. Interactive effects of linguistic abstraction and stimulus statistics in the online modulation of neural speech encoding. Atten Percept Psychophys [Internet]. 2018 Dec 18 [cited 2019 Apr 18]; Available from: https://doi.org/10.3758/s13414-018-1621-9

54. Krishnan A, Gandour JT, Bidelman GM. The effects of tone language experience on pitch processing in the brainstem. Journal of Neurolinguistics. 2010 Jan 1;23(1):81–95.

55. Barron HC, Mars RB, Dupret D, Lerch JP, Sampaio-Baptista C. Cross-species neuroscience: closing the explanatory gap. Philos Trans R Soc Lond B Biol Sci [Internet]. 2021 Jan 4 [cited 2021 Mar 23];376(1815). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7116399/

56. Kriegeskorte N, Mur M, Bandettini PA. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci [Internet]. 2008 [cited 2020 Jul 1];2. Available from: https://www.frontiersin.org/articles/10.3389/neuro.06.004.2008/full

57. Abel TJ, Osorio RV, Amorim-Leite R, Mathieu F, Kahane P, Minotti L, et al. Frameless robot-assisted stereoelectroencephalography in children: technical aspects and comparison with Talairach frame technique. Journal of Neurosurgery: Pediatrics. 2018 Jul 1;22(1):37–46.

58. Krishnan A, Bidelman GM, Gandour JT. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. Hearing Research. 2010 Sep;268(1–2):60–6.

59. Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section. 1994 May 1;92(3):204–14.

60. Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology. 2010 Mar;47(2):236–46.

61. Aiken SJ, Picton TW. Envelope and spectral frequency-following responses to vowel sounds. Hearing Research. 2008 Nov;245(1–2):35–47.

62. Plyler PN, Ananthanarayan AK. Human frequency-following responses: representation of second formant transitions in normal-hearing and hearing-impaired listeners. J Am Acad Audiol. 2001 Dec;12(10):523–33.

63. Wang Q, Li L. Differences between auditory frequency-following responses and onset responses: Intracranial evidence from rat inferior colliculus. Hearing Research. 2018 Jan 1;357:25–32.

64. Du Y, Ma T, Wang Q, Wu X, Li L. Two crossed axonal projections contribute to binaural unmasking of frequency-following responses in rat inferior colliculus. European Journal of Neuroscience. 2009;30(9):1779–89.

65. Coffey EBJ, Musacchia G, Zatorre RJ. Cortical Correlates of the Auditory Frequency-Following and Onset Responses: EEG and fMRI Evidence. The Journal of Neuroscience. 2017 Jan 25;37(4):830–8.

66. Zatorre RJ, Belin P. Spectral and Temporal Processing in Human Auditory Cortex. Cereb Cortex. 2001 Oct 1;11(10):946–53.

67. Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of Phonetic and Pitch Discrimination in Speech Processing. Science. 1992;256(5058):846–9.

68. Lattner S, Meyer ME, Friederici AD. Voice perception: Sex, pitch, and the right hemisphere. Hum Brain Mapp. 2005 Jan;24(1):11–20.

69. Hyde KL, Peretz I, Zatorre RJ. Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia. 2008 Jan 1;46(2):632–9.

70. Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The Processing of Temporal Pitch and Melody Information in Auditory Cortex. Neuron. 2002 Nov 14;36(4):767–76.

71. Ono K, Nakamura A, Yoshiyama K, Kinkori T, Bundo M, Kato T, et al. The effect of musical experience on hemispheric lateralization in musical feature processing. Neuroscience Letters. 2011 Jun 1;496(2):141–5.

72. Krishnan A, Swaminathan J, Gandour JT. Experience-dependent Enhancement of Linguistic Pitch Representation in the Brainstem Is Not Specific to a Speech Context. J Cogn Neurosci. 2009 Jun;21(6):1092–105.

73. Patel AD, Iversen JR. The linguistic benefits of musical abilities. Trends in Cognitive Sciences. 2007 Sep 1;11(9):369–72.

74. Skoe E, Chandrasekaran B. The layering of auditory experiences in driving experience-dependent subcortical plasticity. Hearing Research [Internet]. 2014 Jan [cited 2014 Aug 12]; Available from: http://linkinghub.elsevier.com/retrieve/pii/S0378595514000033

75. Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N. A Toolbox for Representational Similarity Analysis. PLOS Computational Biology. 2014 Apr 17;10(4):e1003553.

76. Peng F, Innes-Brown H, McKay CM, Fallon JB, Zhou Y, Wang X, et al. Temporal Coding of Voice Pitch Contours in Mandarin Tones. Front Neural Circuits [Internet]. 2018 [cited 2021 Apr 27];12. Available from: https://www.frontiersin.org/articles/10.3389/fncir.2018.00055/full

77. Smith JC, Marsh JT, Greenberg S, Brown WS. Human auditory frequency-following responses to a missing fundamental. Science. 1978 Aug 18;201(4356):639–41.

78. Steinschneider M, Volkov IO, Noh MD, Garell PC, Howard MA. Temporal Encoding of the Voice Onset Time Phonetic Parameter by Field Potentials Recorded Directly From Human Auditory Cortex. Journal of Neurophysiology. 1999 Nov 1;82(5):2346–57.

79. Steinschneider M, Fishman YI, Arezzo JC. Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. The Journal of the Acoustical Society of America. 2003 Jul;114(1):307–21.

80. Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. Nature Neuroscience. 2010 Feb;13(2):253–60.

81. Bajo VM, Moore DR. Descending projections from the auditory cortex to the inferior colliculus in the gerbil, Meriones unguiculatus. J Comp Neurol. 2005 May 30;486(2):101–16.

82. Slabu L, Grimm S, Escera C. Novelty Detection in the Human Auditory Brainstem. J Neurosci. 2012 Jan 25;32(4):1447–52.

83. Gnanateja GN, Ranjan R, Firdose H, Sinha SK, Maruthy S. Acoustic basis of context dependent brainstem encoding of speech. Hearing Research. 2013 Oct;304:28–32.

84. Skoe E, Kraus N. Hearing It Again and Again: On-Line Subcortical Plasticity in Humans. Op de Beeck HP, editor. PLoS ONE. 2010 Oct 26;5(10):e13645.

85. Parbery-Clark A, Strait DL, Kraus N. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. Neuropsychologia. 2011 Oct;49(12):3338–45.

86. Petkov CI, Kayser C, Augath M, Logothetis NK. Functional Imaging Reveals Numerous Fields in the Monkey Auditory Cortex. PLOS Biology. 2006 Jun 20;4(7):e215.

87. Wallace MN, Rutkowski RG, Palmer AR. Identification and localisation of auditory areas in guinea pig cortex. Exp Brain Res. 2000 Jun 1;132(4):445–56.

88. Lau JCY, Wong PCM, Chandrasekaran B. Context-dependent plasticity in the subcortical encoding of linguistic pitch patterns. J Neurophysiol. 2017 Feb 1;117(2):594–603.

89. Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, et al. Coding of Repetitive Transients by Auditory Cortex on Heschl's Gyrus. Journal of Neurophysiology. 2009 Aug 12;102(4):2358–74.

90. Naplib/Naplib [Internet]. Naplib; 2020 [cited 2020 Sep 3]. Available from: https://github.com/Naplib/Naplib

91. Faraji AH, Remick M, Abel TJ. Contributions of Robotics to the Safety and Efficacy of Invasive Monitoring With Stereoelectroencephalography. Front Neurol [Internet]. 2020

[cited 2021 Feb 1];11. Available from: https://www.frontiersin.org/articles/10.3389/fneur.2020.570010/full

92. Chabardes S, Abel TJ, Cardinale F, Kahane P. Commentary: Understanding Stereoelectroencephalography: What's Next? Neurosurgery. 2018 Jan 1;82(1):E15–6.

93. Fischl B. FreeSurfer. Neuroimage. 2012 Aug 15;62(2):774–81.

94. Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM. Brainstorm: A User-Friendly Application for MEG/EEG Analysis [Internet]. 2011 [cited 2018 Jun 4]. Available from: https://www.hindawi.com/journals/cin/2011/879716/

95. Teichert T. Tonal frequency affects amplitude but not topography of rhesus monkey cranial EEG components. Hearing Research. 2016 Jun 1;336:29–43.

96. Teichert T, Gurnsey K, Salisbury D, Sweet RA. Contextual processing in unpredictable auditory environments: the limited resource model of auditory refractoriness in the rhesus. Journal of Neurophysiology. 2016 Aug 10;116(5):2125–39.

97. Li F, Teichert T. A surface metric and software toolbox for EEG electrode grids in the macaque. Journal of Neuroscience Methods. 2020 Dec 1;346:108906.

98. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods. 2004;134(1):9–21.

99. Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics. Trends in Cognitive Sciences. 2004 May;8(5):204–10.