Title: Dynamic cortico-cortical information transfer regimes during vocalization.

Authors: Francisco García-Rosales[1*], Luciana López-Jury[1], Eugenia Gonzalez-Palomares[1], Johannes Wetekam[1], Yuranny Cabral-Calderín[2], Ava Kiai[1], Manfred Kössl[1], Julio C. Hechavarría[1*].

Affiliations: [1] Institut für Zellbiologie und Neurowissenschaft, Goethe-Universität, 60438 Frankfurt/M., Germany. [2] Research Group Neural and Environmental Rhythms, Max Planck Institute for Empirical Aesthetics, 60322 Frankfurt/M., Germany.

[*] *Corresponding authors.*

Mailing address:

* Francisco García-Rosales, Institut für Zellbiologie und Neurowissenschaft, Max-von-Laue-Str. 13, 60438 Frankfurt/Main, Germany, Tel.: (+49) 69 / 798 42063. Email: garciarosales@bio.uni-frankfurt.de

* Julio C. Hechavarria, Institut für Zellbiologie und Neurowissenschaft, Max-von-Laue-Str. 13, 60438 Frankfurt/Main, Germany, Tel.: (+49) 69 / 798 42062. Email: hechavarria@bio.uni-frankfurt.de

Abstract

The mammalian frontal and auditory cortices are fundamental structures supporting vocal production, yet the dynamics of information exchange between these regions during vocalization are unknown. Here, we tackle this issue by means of electrophysiological recordings in the fronto-auditory network of freely-vocalizing *Carollia perspicillata* bats. We find that oscillations in frontal and auditory cortices provide correlates of vocal production with complementary patterns across structures. Causality analyses of oscillatory activity revealed directed information exchange in the network, predominantly of top-down nature (frontal to auditory). Such directed connectivity was dynamic, as it depended on the type of vocalization produced, and on the timing relative to vocal onset. Remarkably, we observed the emergence of bottom-up information transfer only when bats produced calls with evident post-vocal consequences (echolocation pulses). Our results link vocal production to dynamic information transfer between sensory (auditory) and association areas in a highly vocal mammalian animal model.

Introduction

Vocal production is a crucial behaviour that underlies the evolutionary success of various animal species. Several cortical and subcortical structures in the mammalian brain support vocalization (Jurgens, 2009), their activities related to vocal control (Gavrilov et al., 2017; Okobi et al., 2019; Zhang and Ghazanfar, 2020), motor preparation (Okobi et al., 2019; Schulz et al., 2005; Tschida et al., 2019), and feedback correction of vocal outputs (Eliades and Tsunada, 2018; Eliades and Wang, 2008). However, the precise neural dynamics that underpin vocal production within these regions, and the nature of long-distance interactions in large-scale neural networks related to vocal utterance, remain poorly understood.

The connectivity patterns of the frontal cortex make it a major hub for cognitive control and behavioural coordination (Choi et al., 2018; Helfrich and Knight, 2019; Zhang et al., 2016). Frontal cortical areas are anatomically connected with structures directly involved in vocal production, such as the periaqueductal grey (Petkov and Jarvis, 2012) and the dorsal striatum (Voorn et al., 2004). Experimental evidence demonstrates that the neural activity in frontal regions relates to vocalization (Gavrilov et al., 2017; Hage and Nieder, 2013; Roy et al., 2016; Weineck et al., 2020), correlating with the acoustic and behavioural properties of produced calls (Hage and Nieder, 2013; Weineck et al., 2020). Frontal regions are also anatomically and functionally connected with the auditory cortex (AC; (García-Rosales et al., 2020; Kobler et al., 1987; Park et al., 2015; Plakke and Romanski, 2014; Winkowski et al., 2013; Winkowski et al., 2018)), a cornerstone structure for audition that exhibits suppression to self-produced sounds (Aliu et al., 2009; Baess et al., 2011; Martikainen et al., 2005; Rummell et al., 2016), including vocalizations (Eliades and Wang, 2003, 2005; Flinker et al., 2010; Tsunada and Eliades, 2020). Such auditory cortical suppression is thought to be mediated by preparatory motor signals originating in the motor system (i.e. “corollary discharges” or “efference copies”; (Clayton et al., 2020; Li et al., 2020)). The attenuation of neural responses in AC during vocal production supports precise vocal control by means of feedback mechanisms (Eliades and Tsunada, 2018; Eliades and Wang, 2008), in which frontal cortical areas are also involved (Behroozmand et al., 2015; Kingyon et al., 2015; Loh et al., 2020; Toyomura et al., 2007). Although current evidence shows that a fronto-auditory cortical circuit is essential for the accurate control of vocal production, the interactions between frontal and auditory cortices during vocalization remain obscure.

In this study, we addressed the neural mechanisms of vocal production in the fronto-auditory network using a highly vocal mammalian model: the bat *Carollia perspicillata* (Fernandez et al., 2014; Hechavarria et al., 2016; Knornschild et al., 2013, 2014). Bats constitute an excellent system to study the underpinnings of vocalization because they rely heavily on vocal behaviour for both communication and navigation. Communication and echolocation calls differ markedly in their spectrotemporal structure (Knornschild et al., 2014) and are vocalized for very different behavioural purposes. The production of these calls is distinctly controlled at the level of the brainstem (Fenzl and Schuller, 2007), possibly mediated by frontal cortical circuits involving regions such as the anterior cingulate cortex (Gooler and O'Neill, 1987) and the frontal-auditory field (FAF; (Weineck et al., 2020)).

Vocal production circuits were studied by measuring local-field potential (LFP) oscillations simultaneously in frontal and auditory cortex regions of vocalizing bats. LFPs are an electrophysiological marker of the extracellular spiking activity and synaptic currents in local neuronal populations (Buzsaki et al., 2012). In frontal and sensory cortices, these signals participate in cognitive processes, sensory computations, and interareal communication via phase coherence (Fries, 2015; García-Rosales et al., 2018; García-Rosales et al., 2020; Helfrich and Knight, 2016; Lakatos et al., 2008; Lakatos et al., 2013). In the FAF, a richly connected auditory region of the bat frontal cortex (Eiermann and Esser, 2000; Kobler et al., 1987), LFP activity predicts vocal output while synchronizing differentially with dorso-striatal oscillations according to vocalization type (Weineck et al., 2020). Neural oscillation in the bat FAF also synchronize across socially interacting bats (Zhang and Yartsev, 2019). In the AC, the roles of oscillatory activity for vocal production are less clear, although human studies suggest that oscillations mediate communication with frontal and motor areas for feedback control (Franken et al., 2018; Kingyon et al., 2015; Schmitt et al., 2020). However, the precise dynamics of information exchange in the fronto-auditory circuit during vocalization are unknown.

We hypothesized the existence of directed information transfer in the FAF-AC network in accordance with both top-down (frontal to auditory) and bottom-up (auditory to frontal) mechanisms for vocal production. The former would be consistent with the roles of frontal regions for vocal coordination; the latter, consistent with the requirements of effective feedback control. By means of simultaneous electrophysiological recordings in the FAF-AC circuit of freely vocalizing bats, we were able to confirm this hypothesis. We report complex causal interactions (within a transfer entropy framework) between frontal and auditory cortices, both during spontaneous activity and periods of vocal production. These interactions were strongly top-down directed. Connectivity patterns were not static, as they varied according to whether animals vocalized echolocation or communication calls and depended on the timing relative to vocal onset. Remarkably, only the production of echolocation pulses resulted in strong and preferential bottom-up information transfer in the auditory-frontal direction after vocalization. Our results suggest that dynamic information transfer in large-scale networks involved in vocal production, such as the FAF-AC circuit, are shaped by the behavioural consequences of produced calls.

Results

Neural activity was studied in the FAF and the AC of *C. perspicillata* bats (3 males) while animals produced self-initiated vocalizations. From a total of 12494 detected vocalizations, 147 echolocation (“sonar”) and 725 non-specific communication (“non-sonar”) calls were preceded by a period of silence lasting at least 500 ms and were therefore considered for subsequent analyses. Representative sonar and non-sonar vocalizations are shown in **Fig. 1a**. Overall, the two types of vocalizations did not differ significantly in terms of call length (Wilcoxon rank sum test, $p = 0.12$; **Fig. 1b**), although call length distributions differed significantly (2-sample Kolmogorov-Smirnov test, $p = 1.48\text{x}10^{-6}$). There were clear differences in the power spectra of sonar and non-sonar calls (**Fig. 1c**, left), such that peak frequencies of sonar utterances were significantly higher than their non-sonar counterparts ($p = 4.48\text{x}10^{-69}$; **Fig. 1c**, right). These spectral differences arise from the stereotypical design of echolocation and communication calls produced by *C. perspicillata* (Hechavarria et al., 2016; Knörnschild et al., 2014).

Oscillations in frontal and auditory cortices predict vocalization type

Figure 1d illustrates electrophysiological activity recorded simultaneously from FAF and AC at various cortical depths, as the sonar and non-sonar vocalizations shown in **Fig. 1a** were produced. Single-trial LFP traces revealed conspicuous pre-vocal oscillatory activity in low and high-frequencies, more pronounced in frontal regions, and strongest when animals produced sonar calls. Power spectral densities (PSD) obtained from pre-vocal LFP segments (i.e. -500 to 0 ms relative to vocal onset; **Fig. 1f**) indicated low- and high-frequency power increase (relative to a no-vocalization baseline, or “no-voc”) associated with vocal production, particularly in FAF and for electrodes located at depths > 100 μm (**Fig. 1e** depicts this at depths of 300 μm; see black arrows). Differences in AC conditional on the type of vocal output were less pronounced and appeared limited to low LFP frequencies (grey arrows in **Fig. 1e**). Such pre-vocal spectral patterns were analysed using canonical LFP frequency bands, namely: delta (δ), 1-4 Hz; theta (θ), 4-8 Hz; alpha (α), 8-12 Hz; low beta (β_1), 12-20 Hz; high beta (β_2), 20-30 Hz; and three sub-bands of gamma (γ): γ_1 (30-60 Hz), γ_2 (60-120 Hz), and γ_3 (120-200 Hz). Pre-vocal LFP power in each band was calculated on a trial-by-trial basis and normalized to no-voc periods.

There were significant power changes between no-voc and pre-vocal periods across frequency bands (**Fig. 1f**, see also **Fig. S1**). Notably, the power increase in low- (δ-α) and high-frequency (γ_2) LFP bands of the FAF was different when animals produced sonar and non-sonar vocalizations, with the highest increase in the pre-vocal sonar case. The opposite pattern was observed in the AC, where differences between ensuing vocalization types were most prominent in β_1 (but not δ-α or γ) frequencies, and were explained by higher pre-vocal power increase for non-sonar than for sonar vocalizations (**Fig. 1f**). Based on these observations, we addressed whether pre-vocal LFP power in frontal and auditory cortices was a significant predictor of ensuing call type. To this effect, generalized linear models (GLMs) were fit using sonar and non-sonar pre-vocal power changes as predictors (see Methods), for all channels (in both structures) and frequency bands. A summary of these models is given in **Fig. 1g** (see the outcomes of two representative GLMs illustrated in **Fig. S1**). Low- and high-frequency power increase (mostly in the δ-α and γ_2 bands) in FAF predicted whether animals produced sonar or non-sonar calls, typically with moderate effect sizes ($p < 0.05$; $R^2m >= 0.1$), highest in middle-to-deep electrodes (i.e. depths > 300 μm; **Fig 1g**, left). In the AC, pre-vocal power predicted ensuing call type mostly in the α-β bands of the spectrum, although more strongly so in β_1 frequencies. Moderate effect sizes were also observed ($p < 0.05$; $R^2m >= 0.1$), which were highest in middle-to-deep electrodes (depths > 350 μm). In summary, these results indicate that pre-vocal oscillatory power significantly predicts ensuing call type in both association (frontal) and sensory (auditory) cortices, though with complementary frequency specificity and opposite effects.

Different morphology of neural oscillations in frontal and auditory cortices

We sought to determine whether the functional differences between frontal and auditory cortical LFPs were echoed by differences in the neural circuitry generating oscillations within each region. The waveform shape of an oscillatory process is a consequence of its underlying neural mechanisms (Cole and Voytek, 2017), and therefore shape differences across LFPs are a proxy of mechanistic differences in their generators. We performed cycle-by-cycle analysis of oscillatory morphology for the LFP activity recorded in FAF and AC (Cole and Voytek, 2019). In the following, we focused on frequency bands that significantly predicted vocal output across structures: δ, θ, α, β_1, and γ_2 (see **Fig. 1g** and **Fig. S1**). For robustness, analyses were performed on whole recordings and not only for LFP segments surrounding vocalizations. Cycles were detected over the raw LFP signal, and only those found in oscillatory bursts were considered (**Fig. 2a** shows examples of detected bursts in δ and γ_2 frequencies). Visual inspection revealed that, for example, average δ- and γ_2- bursts differed between FAF and AC, suggesting not only differences in cycle morphology, but also more "regular" oscillations for FAF LFPs than for those recorded in the AC (**Fig. 2b**, n = 50 bursts). An indicator of the lack of regularity in AC was the "flatter" burst average, which shows that individual burst cycles were more variable (e.g. in terms of period or shape) than those in FAF, and hence more easily averaged-out. Note that, prior to averaging, bursts were normalized and aligned to their second peak. Interestingly, average γ_2 bursts in FAF were embedded in an amplitude dip (**Fig. 2b**, bottom), signalling a relationship between low-frequency phase and high frequency power consistent with previous results in this species (Garcia-Rosales et al., 2020).

Waveform shape was characterized by four main cycle parameters (Cole and Voytek, 2019): rise-decay asymmetry, peak-trough asymmetry, amplitude, and period (**Fig. 2c** illustrates a schematic of their physical meaning). Representative distributions of δ-band period values from ~20 min LFP recordings obtained from FAF and AC at a depth of 700 μm, both recorded simultaneously, are depicted in **Fig. 2d**. Example distributions of other cycle parameters are shown in **Supplementary Figure S2**. While period values in **Fig. 2d** appeared different in frontal and auditory cortices (i.e. higher in frontal areas), another remarkable contrast emerged: the "tightness" of the distributions also differed across structures. Note that the tightness of a cycle parameter distribution indicates the variability of such parameter, and therefore it was used as an indicator of oscillatory shape "regularity" (see above). Distribution tightness was quantified for each channel across penetrations using the Fano factor as a metric (indicated in **Fig. 2d** for the example channels; Fano factor value in FAF: 9.77, in AC: 22.60).

Cycle parameter values and distribution tightness (Fano factor values) were systematically compared between all channel pairs, and across frequency bands. We observed significant differences in parameter values revealing that, indeed, oscillatory morphology differed between FAF and AC (**Supplementary Figure S2**). However, such outcome was not unexpected, as previous work has demonstrated that oscillation shape varies across cortical regions (see (Cole and Voytek, 2017) for a review). What our results indicate is that, besides morphology, oscillatory "regularity" also differs between cortical regions. This was corroborated statistically by comparing Fano factors between areas and recording channels. Plots in **Fig. 2f** show effect size values (Cohen's d) across all pairwise channel comparisons (channels 1-16: FAF, channels 17-32: AC; note the schematic in **Fig. 2e**), with d = 0 for comparisons that were not statistically significant (FDR-corrected Wilcoxon singed-rank tests, significance when $p_{corr} < 0.05$). Channels located in FAF had significantly lower Fano factors across cycle parameters than those located in AC ($p_{corr} < 0.05$; large effect sizes when $|d| > 0.8$, red and blue colours in **Fig. 2f**), mostly for δ, θ, and γ_2 oscillations. For the latter band, however, the effect was the opposite for the parameter amplitude. Conversely, Fano factors from channels in the AC were significantly lower than those in FAF, although only in the β_1-band, for parameters rise-decay asymmetry and period. We noticed that cycles within oscillatory bursts were more regular in frontal or auditory cortices at frequency bands that predicted ensuing vocal type (in FAF: δ, θ, and γ_2; in AC: β_1; see **Fig. 1**). That is, functional differences between FAF and AC were echoed by morphological differences in ongoing oscillations, indicating that a complementary functional link of FAF and AC to vocal production could also be associated to distinct underlying neural mechanisms in each cortical region.

Directed connectivity in the FAF-AC circuit during vocal production

Oscillations in FAF and AC predict ensuing vocal output with functionally opposite patterns, but how rhythms in this network interact during vocal production remains unknown. In previous work we reported low-frequency (1-12 Hz) phase coherence in the FAF-AC circuit during spontaneous activity, with emergence of γ-band (> 25 Hz) coherence at the onset of external acoustic stimulation (García-Rosales et al., 2020). To study FAF-AC oscillatory dynamics during vocal production, we looked beyond phase correlations and examined causal interactions in the fronto-auditory circuit based on a transfer entropy framework. Causal interactions were quantified using directed phase transfer entropy (dPTE), a metric that measures the degree of preferential information transfer between signals based on phase time series (Hillebrand et al., 2016; Lobier et al., 2014). dPTE calculations were performed across vocal conditions for all channel pairs, and for the frequency bands of interest: δ, θ, α, β_1, and γ_2.

Average dPTE connectivity matrices across conditions (sonar and non-sonar pre- and post-vocal periods, and no-voc segments) are illustrated in **Fig. S3**. dPTE matrices were used as adjacency matrices for directed graphs, which characterized patterns of directional information flow in the FAF-AC network (**Fig. 3**). In a graph, nodes represent pooled adjacent channels in either region, according to cortical depth: superficial (*sup*), channels 1-4 (0-150 μm); top-middle (*mid1*), channels 5-8 (200-350 μm); bottom-middle (*mid2*), channels 9-12 (400-550 μm); and *deep*, channels 13-16 (600-750 μm). A directed edge between any two nodes represents preferred information flow between them (e.g. $FAF_{sup} \rightarrow AC_{deep}$). The strength of the directionality was quantified using a directionality index (DI), obtained from normalizing dPTE values to 0.5 (when dPTE = 0.5, there is no preferred direction of information flow). Each edge was weighted according to the DI. The existence of an edge between any two nodes was furthermore conditional on the existence of significant directed connectivity between them based on bootstrap statistics.

During spontaneous activity and pre-vocal periods, significant preferred information flow occurred mostly in the FAF → AC direction, predominantly for δ, θ, and γ_2 frequencies (**Fig. 3a, b**). Connectivity dynamics in these bands indicate that AC oscillatory activity is under top-down influences in both pre-vocal and no-voc periods. Significant FAF → AC preferred directionality of information flow also occurred, albeit more sparsely, in the α and β_1 bands, although the patterns were more variable and differed according to the type of call (sonar vs. non-sonar) produced after the pre-vocal periods (**Fig. 3b**). Preferred information flow occurred in the AC → FAF direction in α (mostly in the pre-vocal non-sonar case) and β_1 (typically for no-voc periods) frequencies. Within-structure directionality of information flow was highest in δ and β_1 bands when considering pre-vocal sonar LFP segments (**Fig. 3b**). Within the FAF, information flow occurred predominantly from deep to superficial layers in δ and β_1 frequencies. Preferential information transfer within FAF was also observed in the α-band, mostly for pre-vocal sonar and no-voc periods, in the superficial-to-deep and deep-to-superficial directions, respectively. In the AC, within-structure information flow was observed for γ_2 frequencies, both during pre-vocal non-sonar and no-voc periods.

Post-vocal directed connectivity patterns were conspicuously different from pre-vocal and spontaneous ones mostly in the δ frequency band (cf. **Fig. 3c** with **Fig. 3a, b**). Whereas, in the pre-vocal sonar case, information flowed mostly in the FAF → AC direction, in the post-vocal sonar case δ-band information flow occurred in the AC → FAF direction. In particular, significant connectivity in the AC → FAF direction occurred in the post-vocal sonar case

(**Fig. 3c**, top) at δ frequencies, originating from the AC_{sup} node (i.e. cortical depths spanning 0-150 μm) and targeting all FAF nodes. Additionally, we observed significant AC → FAF directionality in β_1 frequencies for the post-vocal sonar case, originating from the AC_{mid1}, AC_{mid2}, and AC_{deep} nodes (i.e. depths of 300-750 μm) and targeting all nodes in FAF. Other frequency bands in the post-vocal sonar and non-sonar conditions resembled the existence (or lack) of preferred information flow in the FAF → AC direction observed in the pre-vocal case (**Fig. 3b**). In the frontal cortex, within-structure information flow occurred across frequency bands with various patterns: deep-to-superficial information flow for bands δ (in sonar and non-sonar conditions), α (post-vocal non-sonar), and β_1 (both call types); we also observed superficial-to-deep information flow in the α band for the post-vocal sonar condition. In the AC, within-structure information flow occurred in the deep-to-superficial direction in θ (post-vocal non-sonar), α (both call types), and β_1 (post-vocal sonar) bands; in the superficial-to-deep direction there was information flow in θ (post-vocal sonar; nodes $AC_{top} \rightarrow AC_{mid2}$). The data presented in **Fig. 3** illustrate complex patterns of information exchange within and across the FAF-AC network. Crucially, such patterns vary to a great extent depending on the type of call produced, and on the timing relative to vocal initiation.

Type of vocal output determines connectivity patterns in pre-vocal and post-vocal periods

To quantitatively address the variable information flow shown in **Fig. 3**, we compared connectivity dynamics in the FAF-AC network across vocal conditions (i.e. pre-voc, post-voc, and no-voc), for all the vocalization cases examined.

Connectivity patterns during pre-vocal periods

The top row of **Fig. 4a** summarizes the outcomes of such comparisons during pre-vocal periods across frequency bands, for the sonar vs. non-sonar case. Edges in the graphs are shown if there were significant differences (Wilcoxon rank sum tests, significance when $p < 10^{-4}$) with large effect sizes ($|d| > 0.8$) in the directionality of information flow between two given nodes. Edges were weighted according to the effect size (d) of the corresponding comparisons. Thus, the graphs in **Fig. 4a** (top) show that significant differences (with large effect sizes) between the cases of pre-vocal sonar and pre-vocal non-sonar, in terms of FAF → AC connectivity, occurred only in the γ_2-band. Within-structure directed information flow in the FAF was significantly stronger in the pre-vocal sonar condition when considering LFPs mostly in the δ range. However, sparse significant differences occurred also in the θ and β_1 bands.

Preferred FAF → AC directionality of information flow in the δ band was significantly higher during no-voc periods than during pre-vocal periods related to sonar vocalizations (dashed lines, **Fig. 4b**, top). For γ_2 frequencies, the effect was the opposite: pre-vocal directionality of information flow was significantly higher than that of no-voc periods. Within-structure interactions were strongest in FAF, where the directionality of information flow from bottom to top layers was significantly higher during pre-vocal sonar periods as compared to the other two conditions, in the δ and β_1 bands; the opposite effect was more sparsely seen for α frequencies (**Fig. 4b**, top). Significant differences in the directionality of information flow between non-sonar and no-voc conditions were for the largely inexistent (**Fig. 4c**, top; but note sparse significance in the FAF→AC direction, for the δ-band).

To summarize changes in the directionality of information flow between frontal and auditory cortices, we calculated the net information outflow (DI_{net}) of each area as the sum of the directionality indexes related to outgoing connections from each region. For instance, the DI_{net} of the FAF is the sum of all the edges (i.e. directionality indexes) associated with FAF → AC connections, thus quantifying the net strength of preferential FAF → AC information outflow. Significant differences in the strength of information outflow across conditions (sonar vs. non-sonar, sonar vs. no-voc, and non-sonar vs. no-voc; **Fig. 4a-c**, bottom) occurred with large effect sizes ($|d| > 0.8$) only in the δ and γ_2 bands, when considering information outflow from the FAF. Specifically, FAF-related net information outflow in the γ_2 band was significantly (FDR-corrected Wilcoxon rank sum tests, $p_{corr} < 0.05$) higher when animals vocalized sonar calls as compared to when animals produced non-sonar calls (**Fig. 4a**; $p_{corr} = 1.05\text{x}10^{-83}$, $d = 1.52$) or no call whatsoever (**Fig. 4b**; $p_{corr} = 5.23\text{x}10^{-65}$, $d = 1.26$). Conversely, δ-band net information outflow was significantly higher during no-voc periods as compared to the pre-vocal sonar (**Fig. 4b**; $p_{corr} = 3.37\text{x}10^{-46}$, $d = -1.02$) and, although less prominently, the pre-vocal non-sonar conditions (**Fig. 4c**; $p_{corr} = 1.97\text{x}10^{-26}$, $d = -0.73$).

Connectivity patterns during post-vocal periods

We also observed major differences in connectivity during post-vocal periods between vocalization conditions (**Fig. 5**). Preferential top-down information flow was significantly lower for sonar calls than for non-sonar vocalizations in δ and β_1 frequencies, but significantly higher in the γ_2 band (**Fig. 5a**, top; $p < 10^{-4}$, $|d| > 0.08$). Remarkably, post-vocal preferred directionality of information flow in the δ and β_1 bands was strongest in the bottom-up direction (AC → FAF) for the sonar condition, as opposed to the non-sonar one. Similar effects were seen when comparing connectivity patterns obtained from post-vocal sonar and no-voc periods (**Fig. 5b**, top). In other words, the post-vocal non-sonar condition exhibited the weakest top-down information transfer and the strongest bottom up-information flow in bands δ and β_1. Top-down γ_2 causal influences remained strongest when animals vocalized a sonar call, as compared to non-sonar call production or no-voc periods. Within area changes were observed in the α-band in FAF, where preferential superficial-to-deep information transfer was significantly higher for sonar vocalizations (**Fig. 5a**), while deep-to-superficial information flow was strongest in post-vocal non-sonar and no-voc related periods (Fig. **5b, c**). Finally, significant differences between post-vocal non-sonar and spontaneous activity (**Fig. 5c**, top) were limited to δ frequencies, and strongest for no-voc LFPs.

We compared the net information outflow across conditions in each structure for post-vocal periods (Fig. **5a-c**, bottom). In the δ-band, preferred information outflow from the FAF was weakest (with large effect sizes) when animals vocalized sonar calls (FDR-corrected Wilcoxon rank sum tests; sonar vs. non-sonar: **Fig. 5a**, $p_{corr} = 9.74\text{x}10^{-99}$, d = -1.58; sonar vs. no-voc: **Fig. 5b**, $p_{corr} = 1.90\text{x}10^{-171}$, d = -4.2). A similar effect was observed when comparing non-sonar DI_{net} values with no-voc ones: preferential post-vocal net information outflow from FAF was significantly lower for vocalization-related LFPs (**Fig. 5c**, $p_{corr} = 2.45\text{x}10^{-130}$, d = -2.3). Similarly, post-vocal DI_{net} values for the β_1-band in the FAF were significantly stronger during non-sonar than during sonar vocal production with large effect size (**Fig. 5a**, $p_{corr} = 3.18\text{x}10^{-37}$, d = 0.81). Significant differences in the same frequencies, but between post-vocal sonar and no-voc periods (**Fig. 5b**) did not occur with a large effect size ($p_{corr} = 3.4\text{x}10^{-19}$, d = 0.61). In contrast, γ_2-related net information outflow from FAF was always strongest in the case of sonar vocalizations (sonar vs. non-sonar: **Fig. 6a**, $p_{corr} = 8.89\text{x}10^{-115}$, d = 2.0; sonar vs. no-voc: **Fig. 7b**, $p_{corr} = 7.95\text{x}10^{-90}$, d = 1.59).

The predominance of bottom-up information transfer in low frequencies, dependant on the type of call produced, was evident when considering DI_{net} values. In the δ-band, net information outflow from AC was significantly stronger, with large effect sizes, during sonar production than for post-vocal non-sonar or no-voc periods (sonar vs. non-sonar: **Fig. 5a**, $p_{corr} = 2.68\text{x}10^{-82}$, d = 1.2; sonar vs. no-voc: **Fig. 7b**, $p_{corr} = 1.61\text{x}10^{-124}$, d = 1.49). Also in the β_1-band, net information outflow from AC was strongest for post-vocal sonar than non-sonar periods (**Fig. 5a**; $p_{corr} = 6.31\text{x}10^{-38}$, d = 0.81). However, significant changes between sonar and no-voc cases in the same frequency band did not occur with large effect size (**Fig. 5b**; $p_{corr} = 5.48\text{x}10^{-16}$, d = 0.46). Differences in other frequency bands, or other across-condition comparisons (e.g. non-sonar vs. no-voc, **Fig. 5c**, bottom), were either not reflected in the differential connectivity graphs, or did not have large effect sizes.

Altogether, these results indicate that pre-vocal and post-vocal directional information flow in the FAF-AC network occurs mostly in low and high-frequency bands. The patterns and strength of preferred directionality not only depend on whether a vocalization is produced, but also on the type of vocal output. Crucially, when animals produced non-sonar calls, post-vocal bottom-up influences dominated in δ frequencies, while top-down influences weakened in post-vocal periods compared to spontaneous activity. These results could reflect both a waning of top-down control from the FAF, and an increase in bottom-up transfer in δ and β_1 frequencies. These two possible explanations are not mutually exclusive, and in fact both phenomena may occur in our dataset.

Preferred direction of information flow changes between pre-vocal and post-vocal periods

Differences in the directionality of information flow between pre-vocal and post-vocal activities were addressed by statistically comparing connectivity graphs associated to each case (**Fig. 6**). This is a similar approach to the across-condition comparisons shown in **Figs. 4** and **5**. However, note that paired statistics were performed for these comparisons (Wilcoxon singed-rank tests, significance when $p < 10^{-4}$; see Methods).

In our dataset, FAF → AC preferred information flow was significantly higher (with large effect sizes, $|d| > 0.8$) for pre-vocal periods than for post-vocal ones in the δ and θ bands (**Fig. 6a**, top). For γ_2 frequencies, the effect was the opposite: FAF → AC directionality was highest during post-vocal periods than during pre-vocal ones, with sparse significant differences. Remarkably, AC → FAF preferred directionality of information flow was significantly stronger during post-vocal periods in δ and β_1 frequency bands (**Fig. 6a**). In frontal cortex, differences in within-structure directionality of information flow occurred in frequency bands δ, α, and β_1. In the AC, within structure differences in information flow occurred mostly in α and β_1 bands (although also less consistently in θ and γ_2, **Fig. 6a**), being strongest in the deep-to-superficial direction during post-vocal periods, and in superficial-to-deep directions during pre-vocal periods. Finally, when considering the case of non-sonar call production (**Fig. 6b**, top), differences in the directionality of information flow occurred only in the δ and θ bands, being significantly higher (with large effect sizes) in the FAF → AC direction for pre-vocal periods than for post-vocal ones.

We calculated the net information outflow (DI_{net}) from FAF and AC in order to statistically compare pre-vocal and post-vocal periods in terms of information transfer from each cortical area. Significant differences (FDR-corrected Wilcoxon singed-rank tests, significance for $p_{corr} < 0.05$) with large effect sizes ($|d| > 0.8$) occurred mostly for low and intermediate frequency bands (i.e. δ and β_1) of the LFP. Specifically, for the pre-vocal vs. post-vocal sonar condition (**Fig. 6a**, bottom), the information outflow from FAF was significantly higher in the δ band during pre-vocal periods related to sonar call production ($p_{corr} = 1.87x10^{-82}$, $d = -3.18$). Notably, the net information outflow from AC was significantly higher when considering post-vocal periods than pre-vocal ones ($p_{corr} = 4.04x10^{-63}$, $d = -1.49$). In the β_1 frequency range, there were no significant differences (with large effect sizes) between pre-vocal and post-vocal net information outflow from the FAF. However, DI_{net} values from AC were significantly different with large effect size during post-vocal periods than during pre-vocal ones ($p_{corr} = 3.94x10^{-34}$, $d = -0.87$). Pre-vocal vs. post-vocal comparisons of net information outflow from FAF and AC related to non-sonar vocalizations revealed only significant differences with large effect sizes for δ frequencies in FAF (**Fig. 6b**, bottom). Here, net information outflow was strongest in pre-vocal periods than in post-vocal ones ($p_{corr} = 2.79x10^{-67}$, $d = 1.54$). Other differences related to DI_{net} values occurred but were either not reflected in the differential connectivity graphs (**Fig. 6**, top), or did not have large effect sizes. These results confirm dynamic changes of predominant connectivity patterns in the FAF-AC network from pre-vocal to post-vocal periods, exhibiting frequency specificity and occurring only when animals produce sonar vocalizations.

Discussion

In this study, we addressed the neural dynamics in frontal and auditory cortices during vocal production. Our main findings are as follows (summarized in **Fig. 7**): (i) pre-vocal LFP power in sensory (AC) and association (FAF) cortices predict vocalization type, with LFP frequency specificity and complementary effects across cortical regions; (ii) functional differences between FAF and AC are likely related to distinct neural mechanisms, based on differences on oscillatory morphology; (iii) LFPs in frontal and auditory cortices are causally related (within a TE framework) during vocal production and spontaneous activity; and (iv) connectivity patterns in the FAF-AC network differed across behavioural states (vocalization and spontaneous activity), depended on call type (sonar or non-sonar), and occurred in a frequency specific manner. These findings provide a view on the cortico-cortical network interactions that occur during vocalization in highly vocal mammals.

Pre-vocal LFP power in frontal and auditory cortices predicts ensuing call type

Consistent with previous reports (Gavrilov et al., 2017; Hage and Nieder, 2013; Weineck et al., 2020), our data indicate that neural activity in the frontal cortex predicts vocal output. Thus, oscillations in frontal regions appear instrumental for vocal control. Such position is supported by several lines of evidence, including those below. First, oscillations in the mammalian frontal cortex are involved in cognitive processes and behavioural (also motor) coordination (Gilmartin et al., 2014; Helfrich and Knight, 2016; Pezze et al., 2014). Second, pre-vocal LFP power in frontal areas predicts ensuing call type ((this study, and (Weineck et al., 2020)). Third, low-frequencies in the bat frontal cortex exhibit synchronization patterns with the dorsal striatum (a basal ganglial structure connected to canonical vocal control pathways (Simonyan and Jurgens, 2003)) that are call-type specific (Weineck et al., 2020). Fourth, frontal and auditory oscillatory activities, beyond being phase-synchronized during vocalization (e.g. in humans, (Kingyon et al., 2015)), are causally related with strong top-down influences during pre-vocal periods (current data). We note, however, that the relationship of pre-vocal oscillatory activity and vocalization type shown in this study remains correlational: our data do not establish a causal role of LFPs for the initiation or planning of sonar or non-sonar calls in the bat FAF.

Neural activity in the AC also relates to vocalization (Eliades and Wang, 2003), but the involvement of cortical oscillations in vocal production had so far not been thoroughly examined (see however (Tsunada and Eliades, 2020)). Our results indicate that pre-vocal auditory cortical LFPs, as previously reported with single-unit spiking, relate to vocal initiation. Interestingly, the pre-vocal spectral changes of LFPs in AC were complementary to those seen in the FAF (see **Fig. 1**). Unlike in the FAF, significant pre-vocal power changes in δ-α and γ_2 bands in AC were not call-type specific. Only in the AC, pre-vocal power changes in β_1 predicted whether animals produced sonar or non-sonar calls. While a strongest power increase in FAF signalled the production of a sonar call, higher pre-vocal power in AC was a signature of non-sonar vocalization. Such interesting functional divergences between frontal and auditory regions was accompanied by differences in oscillatory morphology (**Fig. 2**), underscoring the possibility of distinct origins for oscillatory processes within each area.

It is possible to interpret the dynamics of pre-vocal power in AC considering the neural mechanisms related to vocal production in this structure. Neuronal activity in the AC is predominantly suppressed during vocalization, with inhibition at the single neuron level already occurring hundreds of milliseconds prior to call onset (Eliades and Wang, 2003, 2008; Flinker et al., 2010; Forseth et al., 2020). Inhibition in the AC is mediated by motor control regions, which send a copy of the planned motor command to the auditory system (i.e. “corollary discharge” or “efferent copy” signals; (Eliades and Wang, 2013)). A recent study (Li et al., 2020) suggested a distinction between these signals: the first having an overall suppressive effect, independently of the sound being produced; the second carrying specific information about the sound generated, potentially enhancing its future processing. Thus, pre-vocal power changes in low frequencies, undistinguishable across call types, could reflect general inhibitory mechanisms in AC consistent with corollary discharges mediated by higher-order structures. Indeed, our results from causality analyses support the notion of top-down (FAF → AC) control of pre-vocal low-frequency activity. On the other hand, pre-vocal β-band LFPs might constitute oscillatory correlates of efference copies, given the observed call-type specificity (**Fig. 1**). Because FAF → AC causal influences did not equally extend to the β frequencies, pre-vocal β activity in AC might be influenced instead by specialized regions such as the motor cortex, providing a more specific copy of the motor commands required for vocalization. Channels for motor-auditory communication (see (Nelson et al., 2013)) could in fact operate over β frequencies (Abbasi and Gross, 2020; Ford et al., 2008; Franken et al., 2018).

Cycle morphology in FAF and AC: implications of oscillatory regularity

Oscillations in frontal and auditory cortices are not only functionally, but also morphologically different (**Fig. 2**, **S2**). Oscillatory morphology reflects the cellular properties of the generators responsible for recorded mesoscopic rhythms such as LFPs, or EEGs (Cole and Voytek, 2019). In that sense, oscillatory shape differences across cortical areas are likely related to cytoarchitectural differences, and could in fact correlate with the specific functional properties of distinct cortical structures (reviewed in (Cole and Voytek, 2017)). Nevertheless, beyond a direct morphological perspective, our data revealed a remarkable trend: oscillatory regularity differed significantly between frontal and auditory cortices. Differences in regularity (**Fig. 2**) suggest that LFPs in FAF are generated by local networks that oscillate with tighter parameters.

Cycle parameter regularities in FAF and AC provide an interesting perspective on the functional roles of oscillatory processes within each structure. For example, it is possible to speculate that more regular oscillators in FAF could be beneficial for robust interareal communication, which capitalizes on the phase coherence of low and high frequency rhythms (Fries, 2015). Consistent oscillatory activity may act as a reference frame for long-distance interactions, from a central coordinator such as the frontal cortex. This could support cognitive control mechanisms, which rely on low-frequency synchrony between frontal areas and a plethora of brain regions, including sensory cortices (Helfrich and Knight, 2016). Conversely, the AC is a crucial auditory processing structure whose oscillatory activity synchronizes to slow -and fast- rhythms present in external stimuli (García-Rosales et al., 2018; Gross et al., 2013; Kayser et al., 2009; Lakatos et al., 2007; Lakatos et al., 2013; Lakatos et al., 2005; O'Connell et al., 2015). Importantly, oscillations in AC phase-align with external rhythms even when these are not fully periodic (i.e. quasi-periodic), such as speech and natural vocalizations (Giraud and Poeppel, 2012), which requires at least some degree of flexibility (see (Pittman-Polletta et al., 2020)). Less regular oscillators in AC than in FAF could represent a marker of such flexibility, as low-frequency auditory cortical oscillations vary over a wider range of parameters (**Fig. 2**) that could accommodate the variability of the natural rhythms that are to be represented and encoded.

Causal interactions in the FAF-AC network during sonar and non-sonar vocal production

In frontal and auditory cortices, oscillations provide a correlate of vocal production with complementary patterns. In addition, our results uncovered rich causal interactions (within a TE framework) in the FAF-AC circuit with functional relationships to vocalization. In a recent study, we demonstrated that low-frequency FAF-AC coherence occurs even in the absence of acoustic stimulation (i.e. during spontaneous activity; (García-Rosales et al., 2020)). The current results show that interactions in the network go beyond phase correlations, and that during spontaneous activity information flows in low (δ-α) and high (γ_2) frequency bands preferentially from frontal to auditory regions, thus denoting causal top-down influences. Low-frequency top-down influences from higher-order structures (like the FAF) are thought to modulate neuronal activity in sensory cortices according to cognitive variables such as attention, also during spontaneous activity (Fox et al., 2006; Hillebrand et al., 2016; Sang et al., 2017). Attentional modulation from frontal regions facilitates the efficient and selective representation of external stimuli depending on internal behavioural states, which were, however, not explicitly controlled by us during no-voc periods in this study. In general, our data resonate with the hypothesis of spontaneous top-down modulation of oscillatory activity in AC, and suggest a strict control of higher-order structures over sensory areas reflected in concurrent LFP activity across regions.

During pre-vocal periods, we observed changes in the strength of the directionality of information flow related to the vocalization of sonar calls. These changes revealed intriguing transmitter/receiver dynamics in the FAF-AC network that relate to the preparation of a vocal output, and the neural processing of the consequent acoustic inputs such output entails. Consistent with the proposed roles of frontal structures for vocal control, we observed increased within-structure information flow in the FAF prior to vocalization. The dPTE patterns expand the results demonstrating that pre-vocal, frontal LFP power in low- and high-frequencies is a robust correlate of vocal production. Still, it is important to note that vocalization-specific changes in power may affect causality estimations, e.g. by creating confounding differences between the vocal conditions studied. However, the dPTE is a causality estimate that shows robustness to the influence of power, noise, and other variables (Lobier et al., 2018; Young et al., 2017). In our dataset, the pre-vocal δ-band power increase within each region when animals produced sonar vocalizations (call-type specific in FAF, unspecific in AC) was nonetheless accompanied by a decrease of interareal dPTE values. In addition, a δ-band power increase of non-sonar pre-vocal LFPs relative to baseline (**Figs. 1**, **S1**) did not result in significant differences of dPTE values during pre-vocal and spontaneous periods. Thus, changes in causality did not necessarily follow changes in power, as has been reported in previous work (Hillebrand et al., 2016).

Based on the fact that dPTE values related to top-down influences were lowest during pre-vocal sonar periods (**Figs. 3, 4**), it appears that as animals prepare a sonar vocalization, the FAF gradually relinquishes control over the AC in the low-frequency (δ) channel. The weakening of preferred top-down directional information transfer could be taken as a preamble of emerging bottom-up information flow (i.e. in the AC→FAF direction) in δ frequencies after a sonar call is emitted (**Figs. 5, 6**). Remarkably, the same does not happen in the non-sonar case. Echolocation is a vital behaviour for bats, being the predominant strategy for sampling the environment during navigation. After vocalizing a sonar pulse, the bat auditory system must be ready to process incoming echoes and to use this auditory information to construct a representation of surrounding objects (Simmons, 2012), potentially involving higher order structures. The observed switch from top-down to bottom-up processing when animals find themselves in echolocation mode (**Fig. 6**) could in fact represent the readiness of the bat's auditory machinery for the aforementioned task. Concretely, our data suggest that the former may occur over a continuum encompassing a gradual release of the AC from top-down influences (in this case, stemming from the FAF), which in turn opens the way for auditory-frontal information transfer supporting the processing and integration of incoming echoes. In all, processing feedback information directly related to navigation appears to have a larger weight in the bottom-up processing of acoustic cues resulting from a self-generated sound. Echolocation pulses are produced to generate echoes that must be listened to. Communication calls are often targeted to an audience as means of transmitting internal behavioural information (e.g. distress), not aimed at the emitter itself. For the emitter, in such scenario, feedback processing mostly contributes to the adjustment of vocal parameters such as loudness or pitch (Behroozmand et al., 2009; Eliades and Tsunada, 2018; Eliades and Wang, 2012). Since in this study animals vocalized without an audience (i.e. they were isolated in the recording chamber), further research could elucidate whether the presence of conspecifics (i.e. an audience) increases bottom-up information transfer when vocalizing communication calls.

In conclusion, we show that oscillations in frontal and auditory cortices provide a neural correlate of vocal production with remarkable complementary effects across regions. We further demonstrate the existence of complex bi-directional connectivity patterns in the FAF-AC network. The observed top-down influences during pre-vocal periods are consistent with preparatory signals in AC related to vocal initiation which could have frontal or motor cortical provenance. These information flow patterns changed dynamically according to vocalization type and to the timing relative to vocal onset. Crucially, the emergence of strong bottom-up causal influences in the FAF-AC network, only for post-vocal periods associated to sonar call utterance, suggests that the connectivity in the fronto-auditory circuit is shaped by the behavioural implications of the calls produced.

Methods

Animal preparation and surgical procedures

The study was conducted on three awake *Carollia perspicillata* bats (all males). Experimental procedures were in compliance with European regulations for animal experimentation and were approved by the Regierungspräsidium Darmstad (experimental permit #FU-1126). Bats were obtained from a colony at the Goethe University, Frankfurt. Animals used for experiments were kept isolated from the main colony.

Prior to surgical procedures, bats were anaesthetized with a mixture of ketamine (10 mg*kg−1, Ketavet, Pfizer) and xylazine (38 mg*kg−1, Rompun, Bayer). For surgery and for any subsequent handling of the wounds, a local anaesthetic (ropivacaine hydrochloride, 2 mg/ml, Fresenius Kabi, Germany) was applied subcutaneously around the scalp area. A rostro-caudal midline incision was cut, after which muscle and skin tissues were carefully removed in order to expose the skull. A metal rod (ca. 1 cm length, 0.1 cm diameter) was attached to the bone to guarantee head fixation during electrophysiological recordings. The FAF and AC were located by means of well-described landmarks, including the sulcus anterior and prominent blood vessel patterns (see (Eiermann and Esser, 2000; Esser and Eiermann, 1999; García-Rosales et al., 2020)). The cortical surface in these regions was exposed by cutting small holes (ca. 1 mm^2) with the aid of a scalpel blade on the first day of recordings. In the AC, recordings were made mostly in the high frequency fields (Eiermann and Esser, 2000; Esser and Eiermann, 1999; García-Rosales et al., 2020))

After surgery, animals were given no less than two days of rest before the onset of experiments. No experiments on a single animal lasted longer than 4 h per day. Water was given to the bats every 1-1.5 h periods, and experiments were halted for the day if the animal showed any sign of discomfort (e.g. excessive movement). Bats were allowed to rest a full day between consecutive experimental sessions.

Electrophysiological and acoustic recordings

Electrophysiology was performed chronically in fully awake animals, inside a sound-proofed and electrically isolated chamber. Inside the chamber, bats were placed on a custom-made holder which was kept at a constant temperature of 30 °C by means of a heating blanket (Harvard, Homeothermic blanket control unit). Electrophysiological data were acquired from FAF and AC on the left hemisphere, using two 16-channel laminar electrodes (one per structure; Model A1x16, NeuroNexus, MI; 50 μm channel spacing, impedance: 0.5-3 MW per electrode). Probes were carefully inserted into the brain perpendicular to the cortical surface, and lowered with piezo manipulators (one per probe; PM-101, Science 455 products GmbH, Hofheim, Germany) until the top channel was barely visible above the surface of the tissue. The placing and properties of the probes allowed us to record simultaneously at depths ranging from 0-750 μm, spanning all six cortical layers (see (Garcia-Rosales et al., 2019)). Probes were connected to a micro-preamplifier (MPA 16, Multichannel Systems, MCS GmbH, Reutlingen, Germany), and acquisition was done with a single, 32-channel portable system with integrated digitization (sampling frequency, 20 kHz; precision, 16 bits) and amplification steps (Multi Channel Systems MCS GmbH, model ME32 System, Germany).

Acquisition was online-monitored and stored in a computer using the MC_Rack_Software (Multi Channel Systems MCS GmbH, Reutlingen, Germany; version 4.6.2).

Vocal outputs were recorded by means of a microphone (CMPA microphone, Avisoft Bioacustics, Glienicke, Germany) located 10 cm in front of the animal. Recordings were performed with a sampling rate of 250 kHz and a precision of 16 bits. Vocalizations were amplified (gain = 0.5, Avisoft UltraSoundGate 116Hm mobile recording interface system, Glienicke, Germany) and then stored in the same PC used for electrophysiology. Electrophysiological and acoustic data were aligned using two triggers, an acoustic one (5 kHz tone, 10 ms long) presented with a speaker located inside of the chamber (NeoCD 1.0 Ribbon Tweeter; Fountek Electronics), and a TTL pulse sent to the recording system for electrophysiology (see above). Note that the onsets of the tones were in synchrony with the TTL pulses registered by the acquisition system for electrophysiology.

Classification of vocal outputs

Two sessions of concurrent acoustic recordings (~10 min long) were made per paired penetrations in FAF and AC. Vocalizations were automatically detected based on the acoustic envelope of the recordings. The envelope was z-score normalized to a period of no vocalization (no less than 10 s long), which was manually selected, per file, after visual inspection. If a threshold of 5 standard deviations was crossed, a vocalization occurrence was marked and its start and end times were saved. Given the stereotyped spectral properties of *C. perspicillata*'s echolocation calls, a preliminary classification between sonar and non-sonar utterances was done based on each call's peak frequency (a peak frequency > 50 kHz suggested a sonar vocalization, whereas a peak frequency below 50 kHz suggested a non-sonar call). In addition, vocalizations were labelled as candidates for posterior analyses if there was a time of silence no shorter than 500 ms prior to call production to ensure no acoustic contamination on the pre-vocal period that could affect LFP measurements in FAF or AC. Finally, sonar and non-sonar candidate vocalizations were individually and thoroughly examined via visual inspection to validate their classification (sonar or non-sonar), the absence of acoustic contamination in the 500 ms prior to vocal onset, and the correctness of their start and end time stamps. According to the above, and out of a total of 12494 detected vocalizations, 147 sonar and 725 non-sonar calls were then used in further analyses.

Extraction of LFP signals and power analysis

Data analyses were performed using custom-written scripts in MatLab (version 9.5.0.1298439 (R2018b)), Python (version 2.6 or 3.6), and R (RStudio version 1.3.1073). For extracting LFPs, the raw data were band-pass filtered (zero-phase) between 0.1 and 300 Hz (4th order Butterworth filter; *filtfilt* function, MatLab), after which the signals were downsampled to 1 kHz.

All LFP spectral analyses were done using the Chronux toolbox (Bokil et al., 2010). Peri-vocal (i.e. times of -500 - 250 ms relative to vocalization onset) spectrograms (shown in **Fig. 1e**) were obtained using the function *mtspectrumc* with a window of 150 ms, which was slid with 10 ms steps, using 3 tapers with a time-bandwidth product (TW) of 2. Pre-vocal power was calculated with LFP segments spanning -500-0 ms relative to vocal onset, using a TW of 2, and 3 tapers. No-vocalization baseline periods (*no-voc*) with a length of 500 ms were pseudo-randomly selected and their power spectra calculated in order to obtain baseline power values for spontaneous activity. The total number of *no-voc* periods matched the total number of vocalizations (n = 872), in a way that the number of selected *no-voc* periods per recording file matched the number of vocalizations found in that particular file. The power of individual frequency bands (i.e. δ, 1-4 Hz; θ, 4-8 Hz; α, 8-12 Hz; β_1, 12-20 Hz; β_2, 20-30 Hz; γ_1, 30-60 Hz; γ_2, 60-120 Hz; γ_3, 120-200 Hz) was calculated by integration of the power spectral density accordingly for each case. Finally, the increase of pre-vocal power relative to the baseline periods was calculated as follows (per frequency band, on a call-by-call basis):

$$Relative\ power\ change = \frac{BP_{pre-voc} - BP_{no-voc}}{BP_{no-voc}} * 100 \qquad [1],$$

where $BP_{pre\text{-}voc}$ is the pre-vocal power (in the case of either a sonar or non-sonar vocalization) of the given frequency band and a trial (i.e. a specific call), and $BP_{no\text{-}voc}$ is the baseline no-voc power associated to the same frequency band and trial.

Generalized linear model for vocal output prediction

To determine whether pre-vocal power change relative to baseline was able to predict the type of ensuing vocal output, we used a GLM with a logistic link function (i.e. logistic regression). The model analysis was done in Rstudio with the *lme4* package. In brief, logistic regression was used to predict the probability of a binary outcome (0 or 1; non-sonar or sonar, respectively) based on the pre-vocal power change as the predictor variable. The probabilities are mapped by the inverse logit function (sigmoid):

$$\sigma(x) = \frac{1}{1+\exp(-x)} \qquad [2],$$

which restricts the model predictions to the interval [0, 1]. Because of these properties, a logistic regression with GLMs is well suited to compare data (and thus, evaluate predictions of ensuing vocal-output) on a single-trial basis (Zempeltzi et al., 2020).

To estimate the effect size of the fitted models, we used the marginal coefficient of determination (R^2m) with the *MuMIn* pacakage. The R^2m coefficient quantifies the variance in the dependent variable (sonar vs. non-sonar vocalization) explained by the predictor variable (i.e. the relative pre-vocal power change). This value is dimensionless and independent of sample size (Nakagawa and Schielzeth, 2013; Zempeltzi et al., 2020), which makes it ideal to compare effect sizes of different models (e.g. across channels and frequency bands, as in **Fig. 2e**). Effect sizes were considered small when $R^2m < 0.1$, medium when $R^2m >= 0.1$, and large when $R^2m >= 0.4$ (Zempeltzi et al., 2020).

Cycle-by-cycle analysis of oscillations

The evaluation of individual cycle parameters for ongoing oscillations in FAF and AC was done with the *bycycle* package in Python (Cole and Voytek, 2019), and custom-written MatLab scripts for statistical analyses. The *bycycle* package makes possible to detect of individual oscillatory cycles at a given frequency band, and to determine whether such cycles are part of oscillatory bursts (in this study, defined as no less than 3 cycles with stable properties; see below). This approach does not require narrowband-filtering and, by calculating cycle parameters directly on the raw LFP data, avoids methods which rely on sinusoidal basis (such as, for example, Hilbert transforming narrow-band signals).

Burst detection depends on four key parameters that characterize the shape of an individual cycle: rise-decay time asymmetry, peak-to-trough asymmetry, period, and amplitude. A schematic illustrating the meaning of these features is given in **Fig. 3c**. The specific parameters used for the *bycycle* burst detection algorithm are given in **Supplementary Table 1**. Only cycles that were found within detected bursts were considered for further analysis.

Cycle parameters characterize the underlying oscillatory dynamics (Cole and Voytek, 2019), with more tightly distributed parameters for a given LFP signal suggesting more “regular” oscillations. Note that the former does not mean that the oscillation is more or less symmetric, for example, but it does imply a higher consistency of shape. The tightness of a distribution of a parameter (e.g. the period) across cycles was quantified with the Fano factor:

$$F = \frac{\sigma_W^2}{\mu_W} \qquad [3],$$

where σ_W^2 is the variance of the distribution (W), and μ_W its mean.

Fano factors were calculated in FAF and AC, for every channel, frequency band, and cycle parameter. That is, the Fano factor of a channel at a given frequency band and parameter condenses all burst cycles found. It was therefore possible to perform paired statistical comparisons across channels (and thus, also across structures), using signals that were simultaneously recorded (i.e. paired penetrations in FAF and AC, n = 30; FDR-corrected Wilcoxon signed rank tests, significance when $p_{corr} < 0.05$). The cycle parameters themselves were compared across channels in a similar manner. A direct comparison of the parameters does not address oscillatory "regularity" (see above) but allows to determine if two given oscillations have different shapes. For a given channel, penetration, and parameter (e.g. period), the median value of the parameter was obtained. Medians from all penetrations were pairwise compared across channels with paired statistics (n = 30 penetrations; FDR-corrected Wilcoxon signed rank tests, significance when $p_{corr} < 0.05$). All comparisons were performed across parameters and frequency bands.

Directionality analyses

Directional connectivity in the FAF-AC network was quantified with the directed phase transfer entropy (dPTE; (Hillebrand et al., 2016)), based on the phase transfer entropy (PTE) metric (Lobier et al., 2018). PTE is a data-driven, non-parametric directionality index that relates closely to transfer entropy (TE; (Wibral et al., 2014)), but is based on the phase time-series of the signals under consideration (here, FAF and AC field potentials). PTE is sensible to information flow present in broad- and narrowband signals, and is in a large degree robust to the effects of, for example, noise, linear mixing, and sample size (Lobier et al., 2018; Young et al., 2017).

In terms of TE, a signal X causally influences signal Y (both of them can be considered as phase times series), if the uncertainty about the future of Y can be reduced from knowing both the past of signal X and signal Y, as compared to knowing the past of signal Y alone. Formally, the above can be expressed as follows:

$$TE_{xy} = \sum p(Y_{t+\delta}, Y_t, X_t) \log \left(\frac{p(Y_{t+\delta}|Y_t, X_t)}{p(Y_{t+\delta}|Y_t)}\right) \quad [4],$$

where δ represents the delay of the information transfer interaction, and TE_{xy} is the transfer entropy between signals X and Y. The estimation of the probabilities for TE quantification requires large computational times and the tuning of various parameters (Hillebrand et al., 2016). PTE, on the other hand, converts the time series into a sequence of symbols (binned-phase time series, see below), and is able to estimate TE on the phase series reducing significantly both processing times and the necessity for parameter fitting (Lobier et al., 2018).

Phase time series were obtained after filtering the LFP signals in a specific frequency band (e.g. θ, 4-8 Hz) and Hilbert transforming the filtered data. To avoid edge artefacts, the full ~10 minutes recordings were filtered and Hilbert transformed before chunking segments related to individual trials (i.e. pre-voc: -500-0 ms relative to call onset, post-voc: 0-250 ms relative to call onset, or no-voc baseline periods). According to the condition under consideration (sonar/non-sonar and pre-voc/post-voc, or baseline periods), we selected 50 trials pseudo-randomly and then concatenated them before quantifying directional connectivity. This process was repeated 500 times and the distribution of dPTE values obtained from each repetition used for further analyses. The former resulted in a distribution of 500 dPTE connectivity matrices; the median value across these was used for constructing connectivity graphs (see below).

Given the phase of the LFP signals, the PTE was calculated according to equation [4]. However, probabilities in this case were estimated by constructing histograms of binned phases (Lobier et al., 2018) instead of using the full, continuous time series. Following (Scott et al., 1997), the number of bins in the histograms was set to:

$$3.49 * \mu\big(\sigma(\phi)\big) * N_s^{-\frac{1}{3}} \quad [5],$$

where *m* and *s* represent the mean and standard deviation, respectively, *f* represents the phase time series, and N_s denotes the number of samples.

The prediction delay d was set to (N_s x N_{ch})/ N_{+-} (Hillebrand et al., 2016), where N_s and N_{ch} are the number of samples and channels ($N_{ch} = 32$), respectively. The value of N_{+-} corresponds to the number of times the LFP phase changes sign across all channels and times.

The dPTE was calculated from the PTE as follows (Hillebrand et al., 2016):

$$dPTExy = \frac{PTExy}{PTExy + PTEyx} \qquad [6]$$

With values ranging between 0 and 1, dPTEs > 0.5 indicate information flow preferentially in the $X \rightarrow Y$ direction, dPTE values below 0.5 indicate preferential information flow in the opposite direction, and dPTE = 0.5 indicates no preferred direction of information flow. In other words, dPTE is a metric of preferred directionality between two given signals. Note that the dPTE analysis among a set of electrodes yields a directed connectivity matrix that can be considered as an adjacency matrix of a directed graph (see below). All PTE and dPTE calculations were done with the Brainstorm toolbox in MatLab (Tadel et al., 2011).

Connectivity graphs

A graph-theoretic examination of the connectivity patterns was made by constructing directed graphs based on the results obtained from the dPTE analyses (i.e. the median across the 500 repetitions; see above). For simplicity, channels in the FAF and AC within a range of 150 µm were grouped as follows (in the FAF, as an example): FAF_{top}, channels 1-4 (0-150 µm); FAF_{mid1}, channels 5-8 (200-350 µm); FAF_{mid2}, channels 9-12 (400-550 µm); FAF_{bottom}, channels 13-16 (600-750 µm). A similar grouping was done for electrodes located in AC. These channel groups were considered as the nodes of a directed graph. A directed edge (u, v) between any two nodes then represents a preferential information flow from node u to node v. The weight of the edge was taken as the median dPTE for the channel groups corresponding to the nodes, according to the dPTE connectivity matrices. For instance, if the groups considered were FAF_{top} and AC_{bottom}, then the weight between both nodes was the median of the obtained dPTE values calculated from channels 1-4 in FAF towards channels 13-16 in AC. The weight of an edge was quantified as a directionality index (DI):

$$DI = \frac{median(dPTE_{uv}) - 0.5}{0.5} * 100 \qquad [7],$$

which expresses, in percentage points, the strength of the preference of information flow in a certain direction. Equation [6] is based on the fact that a dPTE of 0.5 corresponds to no preferred direction of information flow (Hillebrand et al., 2016).

To statistically validate the directionality shown in the graphs we used a bootstrapping approach. Surrogate adjacency matrices were built for the same channel groups (top, mid1, mid2 and bottom), but electrodes were randomly assigned to each group, independently of their depths or cortical location. This randomization was done independently within each of the 500 dPTE matrices obtained from the main connectivity analysis. Then, an adjacency matrix was obtained from these surrogate data in the same way as described above (i.e. using the median across 500 randomized dPTE matrices). Such a procedure was repeated 10,000 times, yielding an equal number of surrogate graphs. An edge in the original graph was kept if the DI of that edge was at least 2.5 standard deviations higher than the mean of the surrogate distribution obtained for that edge (i.e. higher than the 99.38% of the surrogate observations). Edges that did not fulfill this criterion were labelled as non-significant and were therefore not considered for any subsequent analyses.

Statistical procedures

All statistical analyses were made with custom-written MatLab scripts. Paired and unpaired statistical comparisons were performed with Wilcoxon singed-rank and rank sum tests, respectively. These are appropriately indicated in the text, together with sample sizes and p-values. All statistics, unless otherwise noted, were corrected for multiple comparisons with the False Discovery Rate approach, using the Benjamini and Hochberg procedure (Benjamini and Hochberg, 1995). An alpha of 0.05 was set as threshold for statistical significance. The effect size metric used, unless stated otherwise (as in the GLM case), was Cohen's d:

$$d = \frac{\mu_{D1} - \mu_{D2}}{\sqrt{(\frac{(n_1-1)\sigma_{D1}^2 + (n_2-1)\sigma_{D2}^2}{n_1+n_2-2})}} \qquad [8],$$

where D1 and D2 are two distributions, μ represents the mean, σ^2 represents the variance, while n_1 and n_2 are the sample sizes. Effect sizes were considered small when |d| < 0.5, medium when 0.5 <= |d| <= 0.8, and large when |d| > 0.8 (Cohen, 1988).

To test differences in the connectivity graphs across conditions (e.g. sonar vs. non-sonar), we obtained adjacency matrices for each of the 500 penetrations (one per dPTE connectivity matrix; see above) and compared the distributions using Wilcoxon signed rank tests. Given that the large sample size (n = 500 here) increases the occurrence of significant outcomes in statistical testing, edges were only shown when comparisons were significant and produced large effect sizes (|d| > 0.8).

When comparing connectivity graphs between pre-voc and post-voc conditions, we used the exact same trials per repetitions to construct the distribution of dPTE matrices for the pre- and post-voc cases. A certain repetition *m* for each condition was then treated as paired, and therefore Wilcoxon signed rank tests were used for comparing (as opposed to unpaired statistics above). Again, only edges representing significant differences ($p_{corr} < 0.05$) with large effect sizes were shown.

Conflict of interests

The authors declare no financial or non-financial conflicts of interest.

Acknowledgments

This work was supported by the DFG (Grant No. HE 7478/1-1, to JCH), and the Joachim-Herz Foundation (Fellowship granted to FGR).

References

Abbasi, O., and Gross, J. (2020). Beta-band oscillations play an essential role in motor-auditory interactions. Hum Brain Mapp *41*, 656-665.

Aliu, S.O., Houde, J.F., and Nagarajan, S.S. (2009). Motor-induced suppression of the auditory cortex. J Cogn Neurosci *21*, 791-802.

Baess, P., Horvath, J., Jacobsen, T., and Schroger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. Psychophysiology *48*, 1276-1283.

Behroozmand, R., Karvelis, L., Liu, H., and Larson, C.R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. Clin Neurophysiol *120*, 1303-1312.

Behroozmand, R., Shebek, R., Hansen, D.R., Oya, H., Robin, D.A., Howard, M.A., 3rd, and Greenlee, J.D. (2015). Sensory-motor networks involved in speech production and motor control: an fMRI study. Neuroimage *109*, 418-428.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met *57*, 289-300.

Bokil, H., Andrews, P., Kulkarni, J.E., Mehta, S., and Mitra, P.P. (2010). Chronux: a platform for analyzing neural signals. Journal of neuroscience methods *192*, 146-151.

Buzsaki, G., Anastassiou, C.A., and Koch, C. (2012). The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. Nat Rev Neurosci *13*, 407-420.

Choi, E.Y., Drayna, G.K., and Badre, D. (2018). Evidence for a Functional Hierarchy of Association Networks. J Cogn Neurosci *30*, 722-736.

Clayton, K.K., Williamson, R.S., Hancock, K.E., Tasaka, G.I., Mizrahi, A., Hackett, T.A., and Polley, D.B. (2020). Auditory Corticothalamic Neurons Are Recruited by Motor Preparatory Inputs. Curr Biol.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, 2nd edn (Hillsdale, N.J.: L. Erlbaum Associates).

Cole, S., and Voytek, B. (2019). Cycle-by-cycle analysis of neural oscillations. J Neurophysiol *122*, 849-861.

Cole, S.R., and Voytek, B. (2017). Brain Oscillations and the Importance of Waveform Shape. Trends Cogn Sci *21*, 137-149.

Eiermann, A., and Esser, K.H. (2000). Auditory responses from the frontal cortex in the short-tailed fruit bat Carollia perspicillata. Neuroreport *11*, 421-425.

Eliades, S.J., and Tsunada, J. (2018). Auditory cortical activity drives feedback-dependent vocal control in marmosets. Nature Communications *9*.

Eliades, S.J., and Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J Neurophysiol *89*, 2194-2207.

Eliades, S.J., and Wang, X. (2005). Dynamics of auditory-vocal interaction in monkey auditory cortex. Cereb Cortex *15*, 1510-1523.

Eliades, S.J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature *453*, 1102-1106.

Eliades, S.J., and Wang, X. (2012). Neural correlates of the lombard effect in primate auditory cortex. J Neurosci *32*, 10737-10748.

Eliades, S.J., and Wang, X. (2013). Comparison of auditory-vocal interactions across multiple types of vocalizations in marmoset auditory cortex. J Neurophysiol *109*, 1638-1657.

Esser, K.H., and Eiermann, A. (1999). Tonotopic organization and parcellation of auditory cortex in the FM-bat Carollia perspicillata. Eur J Neurosci *11*, 3669-3682.

Fenzl, T., and Schuller, G. (2007). Dissimilarities in the vocal control over communication and echolocation calls in bats. Behav Brain Res *182*, 173-179.

Fernandez, A.A., Fasel, N., Knörnschild, M., and Richner, H. (2014). When bats are boxing: aggressive behaviour and communication in male Seba's short-tailed fruit bat. Anim Behav *98*, 149-156.

Flinker, A., Chang, E.F., Kirsch, H.E., Barbaro, N.M., Crone, N.E., and Knight, R.T. (2010). Single-trial speech suppression of auditory cortex activity in humans. J Neurosci *30*, 16643-16650.

Ford, J.M., Roach, B.J., Faustman, W.O., and Mathalon, D.H. (2008). Out-of-synch and out-of-sorts: dysfunction of motor-sensory communication in schizophrenia. Biol Psychiatry *63*, 736-743.

Forseth, K.J., Hickok, G., Rollo, P.S., and Tandon, N. (2020). Language prediction mechanisms in human auditory cortex. Nat Commun *11*, 5240.

Fox, M.D., Corbetta, M., Snyder, A.Z., Vincent, J.L., and Raichle, M.E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proc Natl Acad Sci U S A *103*, 10046-10051.

Franken, M.K., Eisner, F., Acheson, D.J., McQueen, J.M., Hagoort, P., and Schoffelen, J.M. (2018). Self-monitoring in the cerebral cortex: Neural responses to small pitch shifts in auditory feedback during speech production. Neuroimage *179*, 326-336.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. Neuron *88*, 220-235.

García-Rosales, F., Beetz, M.J., Cabral-Calderin, Y., Kössl, M., and Hechavarria, J.C. (2018). Neuronal coding of multiscale temporal features in communication sequences within the bat auditory cortex. Communications Biology *1*, 200.

García-Rosales, F., López-Jury, L., González-Palomares, E., Cabral-Calderín, Y., and Hechavarría, J.C. (2020). Fronto-Temporal Coupling Dynamics During Spontaneous Activity and Auditory Processing in the Bat Carollia perspicillata. Frontiers in Systems Neuroscience *14*.

Garcia-Rosales, F., Lopez-Jury, L., Gonzalez-Palomares, E., Cabral-Calderin, Y., Kossl, M., and Hechavarria, J.C. (2020). Phase-amplitude coupling profiles differ in frontal and auditory cortices of bats. Eur J Neurosci.

Garcia-Rosales, F., Rohrig, D., Weineck, K., Rohm, M., Lin, Y.H., Cabral-Calderin, Y., Kossl, M., and Hechavarria, J.C. (2019). Laminar specificity of oscillatory coherence in the auditory cortex. Brain Struct Funct *224*, 2907-2924.

Gavrilov, N., Hage, S.R., and Nieder, A. (2017). Functional Specialization of the Primate Frontal Lobe during Cognitive Control of Vocalizations. Cell Rep *21*, 2393-2406.

Gilmartin, M.R., Balderston, N.L., and Helmstetter, F.J. (2014). Prefrontal cortical regulation of fear learning. Trends Neurosci *37*, 455-464.

Giraud, A.L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci *15*, 511-517.

Gooler, D.M., and O'Neill, W.E. (1987). Topographic representation of vocal frequency demonstrated by microstimulation of anterior cingulate cortex in the echolocating bat, Pteronotus parnelli parnelli. J Comp Physiol A *161*, 283-294.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., and Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. PLoS Biol *11*, e1001752.

Hage, S.R., and Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. Nat Commun *4*, 2409.

Hechavarria, J.C., Beetz, M.J., Macias, S., and Kossl, M. (2016). Distress vocalization sequences broadcasted by bats carry redundant information. J Comp Physiol A Neuroethol Sens Neural Behav Physiol *202*, 503-515.

Helfrich, R.F., and Knight, R.T. (2016). Oscillatory Dynamics of Prefrontal Cognitive Control. Trends Cogn Sci *20*, 916-930.

Helfrich, R.F., and Knight, R.T. (2019). Cognitive neurophysiology of the prefrontal cortex. Handb Clin Neurol *163*, 35-59.

Hillebrand, A., Tewarie, P., van Dellen, E., Yu, M., Carbo, E.W., Douw, L., Gouw, A.A., van Straaten, E.C., and Stam, C.J. (2016). Direction of information flow in large-scale resting-state networks is frequency-dependent. Proc Natl Acad Sci U S A *113*, 3867-3872.

Jurgens, U. (2009). The neural control of vocalization in mammals: a review. J Voice *23*, 1-10.

Kayser, C., Montemurro, M.A., Logothetis, N.K., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron *61*, 597-608.

Kingyon, J., Behroozmand, R., Kelley, R., Oya, H., Kawasaki, H., Narayanan, N.S., and Greenlee, J.D. (2015). High-gamma band fronto-temporal coherence as a measure of functional connectivity in speech motor control. Neuroscience *305*, 15-25.

Knornschild, M., Feifel, M., and Kalko, E.K.V. (2013). Mother-offspring recognition in the bat Carollia perspicillata. Anim Behav *86*, 941-948.

Knornschild, M., Feifel, M., and Kalko, E.K.V. (2014). Male courtship displays and vocal communication in the polygynous bat Carollia perspicillata. Behaviour *151*, 781-798.

Knörnschild, M., Feifel, M., and Kalko, E.K.V. (2014). Male courtship displays and vocal communication in the polygynous bat Carollia perspicillata. Behaviour *151*, 781-798.

Kobler, J.B., Isbey, S.F., and Casseday, J.H. (1987). Auditory pathways to the frontal cortex of the mustache bat, Pteronotus parnellii. Science *236*, 824-826.

Lakatos, P., Chen, C.M., O'Connell, M.N., Mills, A., and Schroeder, C.E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron *53*, 279-292.

Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science *320*, 110-113.

Lakatos, P., Musacchia, G., O'Connel, M.N., Falchier, A.Y., Javitt, D.C., and Schroeder, C.E. (2013). The Spectrotemporal Filter Mechanism of Auditory Selective Attention. Neuron *77*, 750-761.

Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., and Schroeder, C.E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol *94*, 1904-1911.

Li, S., Zhu, H., and Tian, X. (2020). Corollary Discharge Versus Efference Copy: Distinct Neural Signals in Speech Preparation Differentially Modulate Auditory Responses. Cereb Cortex *30*, 5806-5820.

Lobier, M., Palva, J.M., and Palva, S. (2018). High-alpha band synchronization across frontal, parietal and visual cortex mediates behavioral and neuronal effects of visuospatial attention. Neuroimage *165*, 222-237.

Lobier, M., Siebenhuhner, F., Palva, S., and Palva, J.M. (2014). Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. Neuroimage *85 Pt 2*, 853-872.

Loh, K.K., Procyk, E., Neveu, R., Lamberton, F., Hopkins, W.D., Petrides, M., and Amiez, C. (2020). Cognitive control of orofacial motor and vocal responses in the ventrolateral and dorsomedial human frontal cortex. Proc Natl Acad Sci U S A *117*, 4994-5005.

Martikainen, M.H., Kaneko, K., and Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. Cereb Cortex *15*, 299-302.

Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol *4*, 133-142.

Nelson, A., Schneider, D.M., Takatoh, J., Sakurai, K., Wang, F., and Mooney, R. (2013). A circuit for motor cortical modulation of auditory cortical activity. J Neurosci *33*, 14342-14353.

O'Connell, M.N., Barczak, A., Ross, D., McGinnis, T., Schroeder, C.E., and Lakatos, P. (2015). Multi-Scale Entrainment of Coupled Neuronal Oscillations in Primary Auditory Cortex. Front Hum Neurosci *9*, 655.

Okobi, D.E., Jr., Banerjee, A., Matheson, A.M.M., Phelps, S.M., and Long, M.A. (2019). Motor cortical control of vocal interaction in neotropical singing mice. Science *363*, 983-988.

Park, H., Ince, R.A., Schyns, P.G., Thut, G., and Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. Curr Biol *25*, 1649-1653.

Petkov, C.I., and Jarvis, E.D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. Front Evol Neurosci *4*, 12.

Pezze, M., McGarrity, S., Mason, R., Fone, K.C., and Bast, T. (2014). Too little and too much: hypoactivation and disinhibition of medial prefrontal cortex cause attentional deficits. J Neurosci *34*, 7931-7946.

Pittman-Polletta, B., Wang, Y., Stanley, D.A., Schroeder, C., Whittington, M., and Kopell, N. (2020). Differential contributions of synaptic and intrinsic inhibitory currents to speech segmentation via flexible phase-locking in neural oscillators. bioRxiv.

Plakke, B., and Romanski, L.M. (2014). Auditory connections and functions of prefrontal cortex. Front Neurosci-Switz *8*.

Roy, S., Zhao, L., and Wang, X. (2016). Distinct Neural Activities in Premotor Cortex during Natural Vocal Behaviors in a New World Primate, the Common Marmoset (Callithrix jacchus). J Neurosci *36*, 12168-12179.

Rummell, B.P., Klee, J.L., and Sigurdsson, T. (2016). Attenuation of Responses to Self-Generated Sounds in Auditory Cortical Neurons. J Neurosci *36*, 12010-12026.

Sang, N., Zhang, L., Hao, L., Wang, Y., Wang, X., Zhang, F., Huang, H., Hou, X., Mao, Y., Bi, T.*, et al.* (2017). Human sensory cortex structure and top-down controlling brain network determine individual differences in perceptual alternations. Neurosci Lett *636*, 113-119.

Schmitt, L.M., Wang, J., Pedapati, E.V., Thurman, A.J., Abbeduto, L., Erickson, C.A., and Sweeney, J.A. (2020). A neurophysiological model of speech production deficits in fragile X syndrome. Brain Commun *2*.

Schulz, G.M., Varga, M., Jeffires, K., Ludlow, C.L., and Braun, A.R. (2005). Functional neuroanatomy of human vocalization: an H215O PET study. Cereb Cortex *15*, 1835-1847.

Scott, S.K., Young, A.W., Calder, A.J., Hellawell, D.J., Aggleton, J.P., and Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. Nature *385*, 254-257.

Simmons, J.A. (2012). Bats use a neuronally implemented computational acoustic model to form sonar images. Curr Opin Neurobiol *22*, 311-319.

Simonyan, K., and Jurgens, U. (2003). Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. Brain Res *974*, 43-59.

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., and Leahy, R.M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. Comput Intell Neurosci *2011*, 879716.

Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., and Kuriki, S. (2007). Neural correlates of auditory feedback control in human. Neuroscience *146*, 499-503.

Tschida, K., Michael, V., Takatoh, J., Han, B.X., Zhao, S., Sakurai, K., Mooney, R., and Wang, F. (2019). A Specialized Neural Circuit Gates Social Vocalizations in the Mouse. Neuron *103*, 459-472 e454.

Tsunada, J., and Eliades, S.J. (2020). Dissociation of Unit Activity and Gamma Oscillations during Vocalization in Primate Auditory Cortex. J Neurosci *40*, 4158-4171.

Voorn, P., Vanderschuren, L.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci *27*, 468-474.

Weineck, K., Garcia-Rosales, F., and Hechavarria, J.C. (2020). Neural oscillations in the fronto-striatal network predict vocal output in bats. PLoS Biol *18*, e3000658.

Wibral, M., Vicente, R., and Lindner, M. (2014). Transfer Entropy in Neuroscience. In Directed Information Measures in Neuroscience, M. Wibral, R. Vicente, and J.T. Lizier, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 3-36.

Winkowski, D.E., Bandyopadhyay, S., Shamma, S.A., and Kanold, P.O. (2013). Frontal cortex activation causes rapid plasticity of auditory cortical processing. J Neurosci *33*, 18134-18148.

Winkowski, D.E., Nagode, D.A., Donaldson, K.J., Yin, P., Shamma, S.A., Fritz, J.B., and Kanold, P.O. (2018). Orbitofrontal Cortex Neurons Respond to Sound and Activate Primary Auditory Cortex Neurons. Cereb Cortex *28*, 868-879.

Young, C.K., Ruan, M., and McNaughton, N. (2017). A Critical Assessment of Directed Connectivity Estimates with Artificially Imposed Causality in the Supramammillary-Septo-Hippocampal Circuit. Front Syst Neurosci *11*, 72.

Zempeltzi, M.M., Kisse, M., Brunk, M.G.K., Glemser, C., Aksit, S., Deane, K.E., Maurya, S., Schneider, L., Ohl, F.W., Deliano, M., *et al.* (2020). Task rule and choice are reflected by layer-specific processing in rodent auditory cortical microcircuits. Commun Biol *3*, 345.

Zhang, S., Xu, M., Chang, W.C., Ma, C., Hoang Do, J.P., Jeong, D., Lei, T., Fan, J.L., and Dan, Y. (2016). Organization of long-range inputs and outputs of frontal cortex for top-down control. Nat Neurosci *19*, 1733-1742.

Zhang, W., and Yartsev, M.M. (2019). Correlated Neural Activity across the Brains of Socially Interacting Bats. Cell *178*, 413-428 e422.

Zhang, Y.S., and Ghazanfar, A.A. (2020). A Hierarchy of Autonomous Systems for Vocal Production. Trends Neurosci *43*, 115-126.

Figure legends

Fig. 1. Pre-vocal oscillations in frontal and auditory cortices allow to predict ensuing vocal output. (**a**) Oscillograms (top) and spectrogram (bottoms) of exemplary sonar (left) and non-sonar calls produced by *C. perspicillata*. (**b**) Distribution of sonar (blue, n = 147) and non-sonar (orange, n = 725) call lengths. No significance differences were observed (Wilcoxon rank sum test, p = 0.12). (**c**) (*Left*) Normalized average power spectral density (PSD) of sonar (blue) and non-sonar (orange) calls. (*Right*) Distribution of peak frequencies of sonar and non-sonar utterances. There were significance differences in the peak frequencies of sonar and non-sonar vocalizations ($p = 4.48x10^{-69}$). (**d**) Single-trial LFPs recorded simultaneous to sonar and non-sonar utterances (left and right columns, respectively). The vertical red dashed line, at time 0, indicates the moment of vocalization onset. The top 16 traces correspond to LFPs recorded in the FAF; the bottom 16 LFP traces were recorded from the AC (demarcated in the panel). FAF and AC recordings were made simultaneously. (**e**) Average pre-vocal (-500 to 0 ms, relative to call onset) power spectral densities (PSD) at a representative depth (300 μm) in FAF and AC. Blue: sonar calls; orange: non-sonar calls; black, dashed: no-voc periods. The difference between sonar and

non-sonar PSDs is depicted in grey (right). (**e**) Percentage pre-vocal power change across representative LFP bands (δ, 1-4 Hz; β_1, 12-20 Hz; γ_2, 60-120 Hz), relative to a no-voc baseline, across cortical depths in FAF (top) and AC (bottom). Values related to sonar utterances (n = 147) are depicted in blue; those related to non-sonar utterances (n = 725) are depicted in orange. Data shown as mean ± sem. (**f**) Pre-vocal power change in frontal and auditory regions are predictors of vocalization type. Effect size (R^2_m) of GLMs considering all frequency bands and channels, both in frontal and auditory cortices. Effect sizes were considered small when $R^2_m < 0.1$, and medium for $R^2_m >= 0.1$. For illustrative purposes, effect size values from non-significant models were set to 0.

Fig. 2. Cycle- parameter regularity differs between frontal and auditory cortices. (**a**) Simultaneously recorded, 30s long raw LFP traces (grey) from FAF and AC at depths of 700 μm, in which detected oscillatory bursts in δ and γ_2 bands are marked in black and magenta, respectively. (**b**) Average δ and γ_2 bursts from frontal and auditory cortices (n = 50 bursts). (**c**) Schematic representation of cycle parameters used to characterize oscillatory morphology. (**d**) Representative distribution of cycle parameter "period" for δ LFPs in simultaneously recorded FAF (blue) and AC (orange) channels (same penetration in **a**). Fano factor in FAF: 9.77; Fano factor in AC: 22.60. (**e**) Schematic representation illustrating the association between region and cortical depth with the channel number identifier of panel **f**. (**f**) Effect sizes (Cohen's *d*) of the channel-by-channel pairwise comparisons between cycle parameters, across penetrations and frequency bands (δ, θ, α, β_1, γ_2). In each plot, a pixel with index *(i, j)* depicts the *d* obtained from comparing channels *i* and *j* (solid black lines in the plots divide channels from frontal and auditory regions). *d* values are only shown if the statistical comparisons are significant (FDR-corrected Wilcoxon signed rank tests, $p_{corr} < 0.05$); the value was set to 0 otherwise ($p_{corr} >= 0.05$). Note that if the pixel with indices *i, j* is coloured blue, then channel with index *i* had a significantly tighter distribution (lower Fano factor) than the channel with index *j*. Large effect sizes occur for |d| > 0.8 (notably, blue and red regions in the plots).

Fig. 3. Directed connectivity patterns in the FAF-AC network. (**a**) Graph visualization of directed connectivity in the FAF network during no-voc periods. FAF and AC channels were condensed into four categories: top (0-150 μm), mid1 (200-350 μm), mid2 (400-550 μm), and bottom (600-750 μm). Graph edges are weighted according to the strength of the preferred directionality (FAF->AC in blue; AC->FAF in orange; within area directionality in grey). Edges are only shown if there was significant preferred directionality according to a threshold defined by bootstrapping. (**b**) Similar to **a**, but directed connectivity was calculated in the pre-vocal sonar and non-sonar conditions. (**c**) Same as **b**, with connectivity patterns obtained for post-vocal sonar and non-sonar conditions.

Fig. 4. Pre-vocal dPTE differences across vocalization conditions. (**a**) (*Top*) Graphs illustrating the differences in pre-vocal directionality between vocalization conditions (sonar vs. non-sonar), across frequency bands. Edges were shown if three conditions were met: (i) the differences were significant (FDR-corrected Wilcoxon rank sum tests, $p_{corr} <$ 0.05), (ii) the effect size was large (|*d*| > 0.8), and (iii) edges had already shown significant directionality (see edges in **Fig. 3**). Edge thickness is weighted according to the effect size

of the comparison. Continuous lines indicate dPTEs for sonar (first condition) higher than dPTEs for non-sonar (second condition) call production. Dashed lines indicate the opposite. (*Bottom*) Net information outflow (DI_{net}) from FAF (blue bars) and AC (orange bars), in the two conditions considered (sonar vs. non-sonar vocal production). Significant differences across conditions are marked with stars (FDR-corrected Wilcoxon signed-rank tests; * $p_{corr} < 0.05$, ** $p_{corr} < 0.01$, *** $p_{corr} < 0.001$, n.s.: not significant; n = 500). Grey numbers in the panels indicate effect sizes (d; not shown for non-significant differences). Here, values are considered independently of whether there was previous significant directionality in any of the two conditions. Data shown as mean ± sem. (**b**) Same as in **a**, but comparing between sonar vs. no-voc conditions. (**c**) Same as in **a** and **b**, but comparing between non-sonar and no-voc conditions.

Fig. 5. Post-vocal dPTE differences across vocalization conditions. (**a**) (*Top*) Graphs illustrating the differences in post-vocal directionality between vocalization conditions (sonar vs. non-sonar), across frequency bands. Edges were shown if three conditions were met: (i) the differences were significant (FDR-corrected Wilcoxon rank sum tests, $p_{corr} < 0.05$), (ii) the effect size was large ($|d| > 0.8$), and (iii) edges had already shown significant directionality (see edges in **Fig. 3**). Edge thickness is weighted according to the effect size of the comparison. Continuous lines indicate dPTEs for sonar (first condition) higher than dPTEs for non-sonar (second condition) call production. Dashed lines indicate the opposite. (*Bottom*) Net information outflow (DI_{net}) from FAF (blue bars) and AC (orange bars), in the two conditions considered (sonar vs. non-sonar vocal production). Significant differences across conditions are marked with stars (FDR-corrected Wilcoxon signed-rank tests; * $p_{corr} < 0.05$, ** $p_{corr} < 0.01$, *** $p_{corr} < 0.001$, n.s.: not significant; n = 500). Grey numbers in the panels indicate effect sizes (d; not shown for non-significant differences). Here, values are considered independently of whether there was previous significant directionality in any of the two conditions. Data shown as mean ± sem. (**b**) Same as in **a**, but comparing between sonar vs. no-voc conditions. (**c**) Same as in **a** and **b**, but comparing between non-sonar and no-voc conditions.

Fig. 6. Pre-vocal and post-vocal directionality differences in the FAF-AC network. (**a**) (*Top*) Graphs illustrating the differences between pre-vocal and post-vocal directionality, across frequency bands and during the production of sonar calls. Edges were shown if two conditions were met: (i) the differences were significant (FDR-corrected Wilcoxon rank sum tests, $p_{corr} < 0.05$), (ii) the effect size was large ($|d| > 0.8$), and (iii) edges had already shown significant directionality (see edges in **Fig. 3**). Edge thickness is weighted according to the effect size of the comparison. Continuous lines indicate pre-vocal dPTEs (first condition) higher than post-vocal dPTEs. Dashed lines indicate the opposite. (*Bottom*) Net information outflow (DI_{net}) from FAF (blue bars) and AC (orange bars), in the two conditions considered (pre-vocal vs. post-vocal). Significant differences between conditions are marked with stars (FDR-corrected Wilcoxon signed-rank tests; * $p_{corr} < 0.05$, ** $p_{corr} < 0.01$, *** $p_{corr} < 0.001$, n.s.: not significant; n = 500). Grey numbers in the panels indicate effect sizes (d; not shown for non-significant differences). Here, values are considered independently of whether there was previous significant directionality in any of the two conditions. Data shown as mean ± sem. (**b**) Same as in **a**, but illustrating

comparisons of directionality between pre-vocal vs. post-vocal conditions related to the vocalization of non-sonar calls.

Fig. 7. The FAF-AC network during vocal production. (**a**) A schematic of *C. perspicillata*'s brain, illustrating the location of the FAF (blue) and the AC (orange). LFPs in these two structures differ in their cycle morphology and, remarkably, in their cycle parameter regularity. Oscillations in frontal and auditory cortices provide a neural correlate vocal production, allowing the prediction of ensuing call type. Prediction is possible in complementary frequency bands in each region, and with opposite effects. (**b**) Schematic representation of causal interactions (within a TE framework) in the FAF-AC network. Strong top-down control, mostly in δ and γ frequencies, occurs during spontaneous activity (no-voc) and prior to vocal utterance. In the δ-band, information flows top-down in the circuit (FAF → AC) during pre-vocal periods, but changes to bottom-up (AC → FAF) information transfer during post-vocal periods. The directionality patterns and the strength of preferential causal interactions depend on the type of call produced, and on the timing relative to vocal onset.

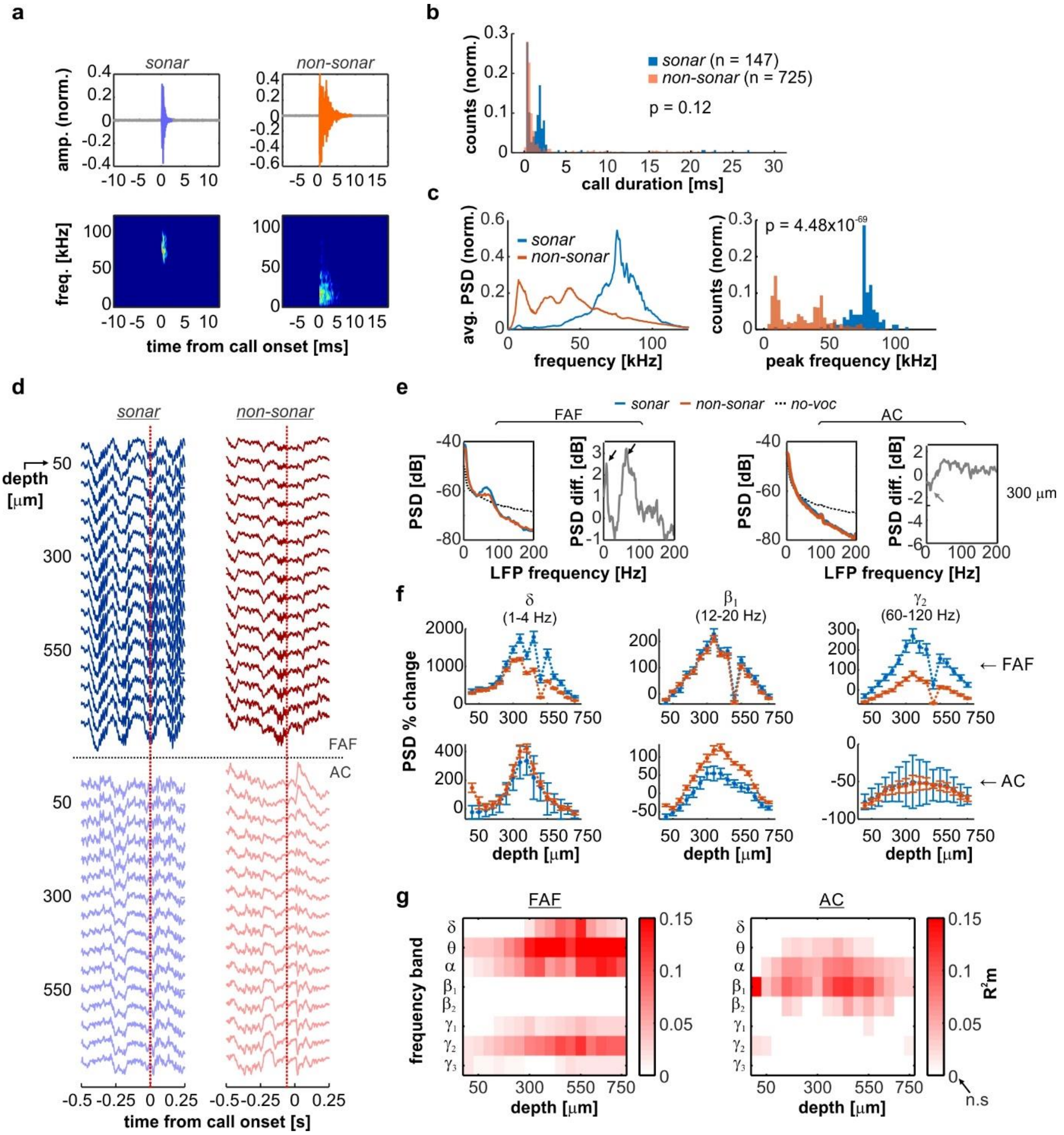


Fig. 1. Pre-vocal oscillations in frontal and auditory cortices allow to predict ensuing vocal output.

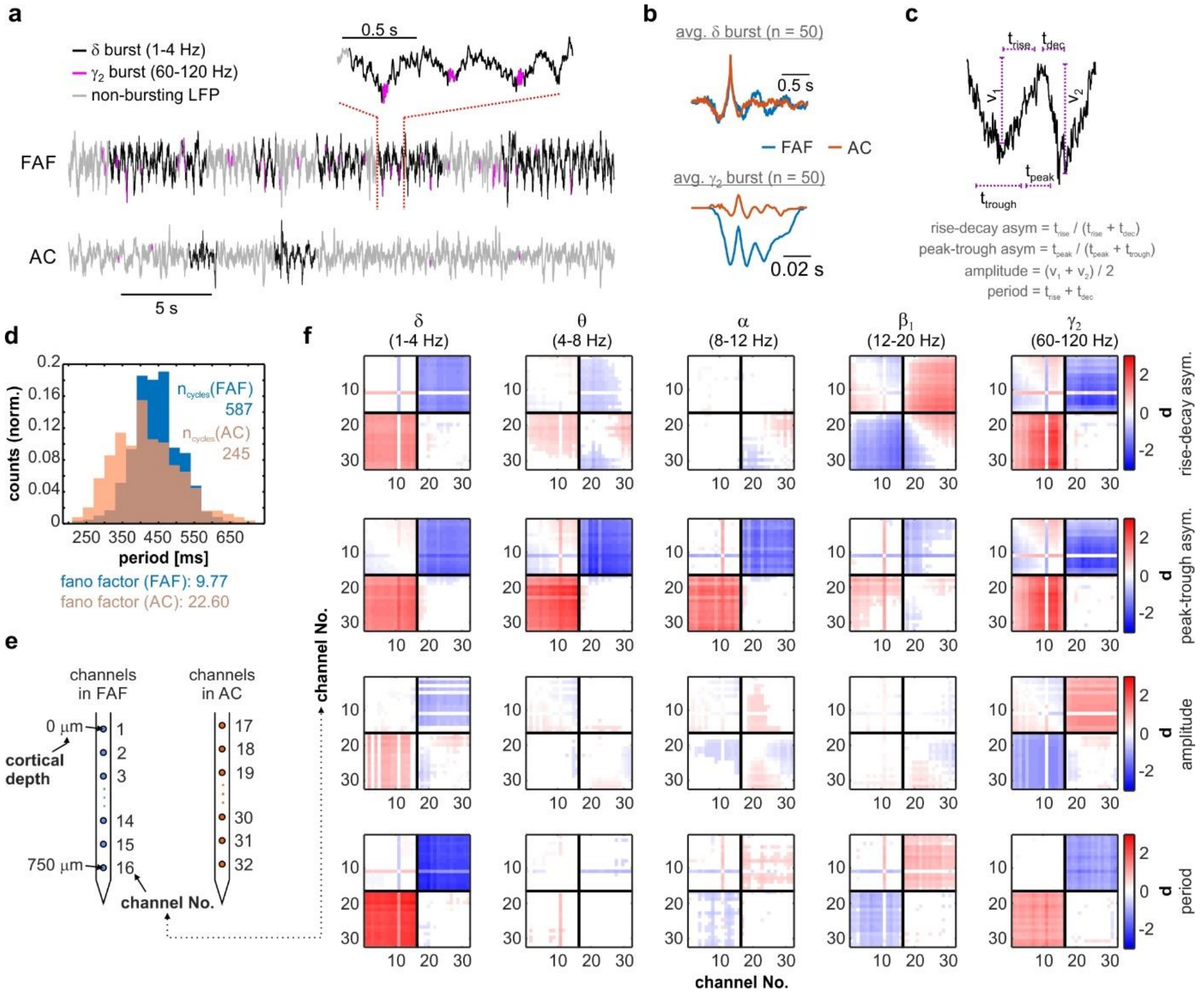


Fig. 2. Cycle- parameter regularity differs between frontal and auditory cortices.

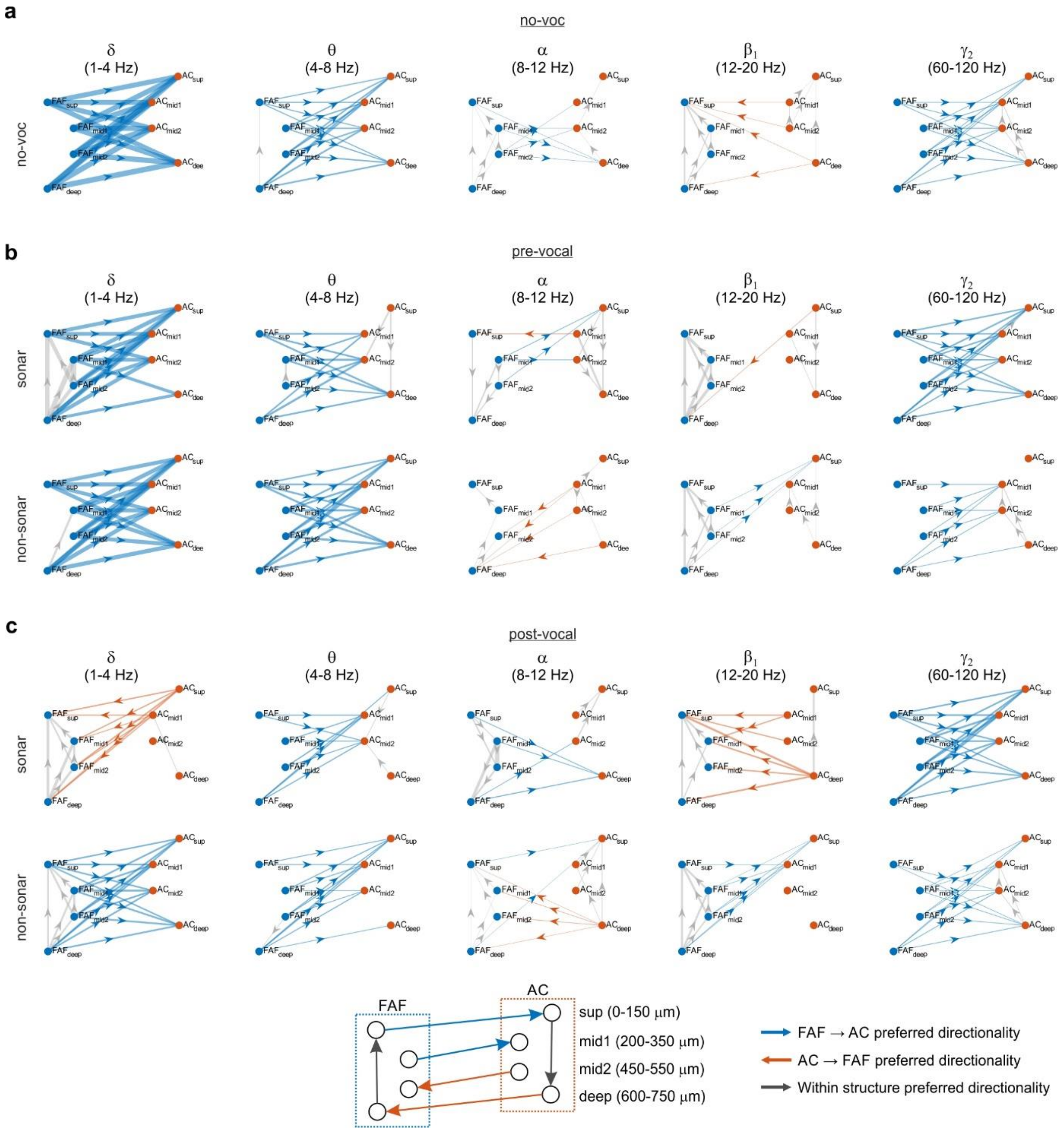


Fig. 3. Directed connectivity patterns in the FAF-AC network.

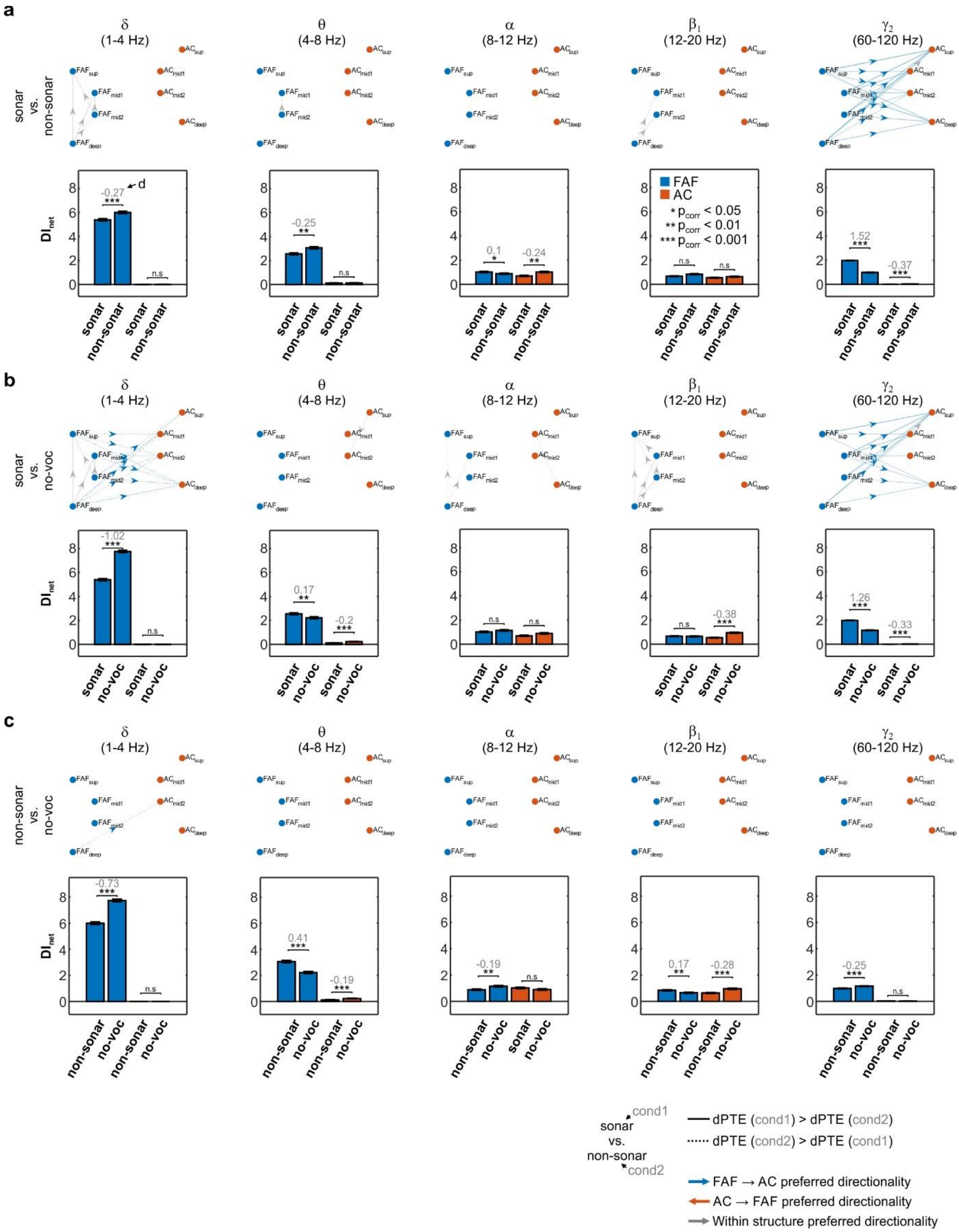


Fig. 4. Pre-vocal dPTE differences across vocalization conditions.

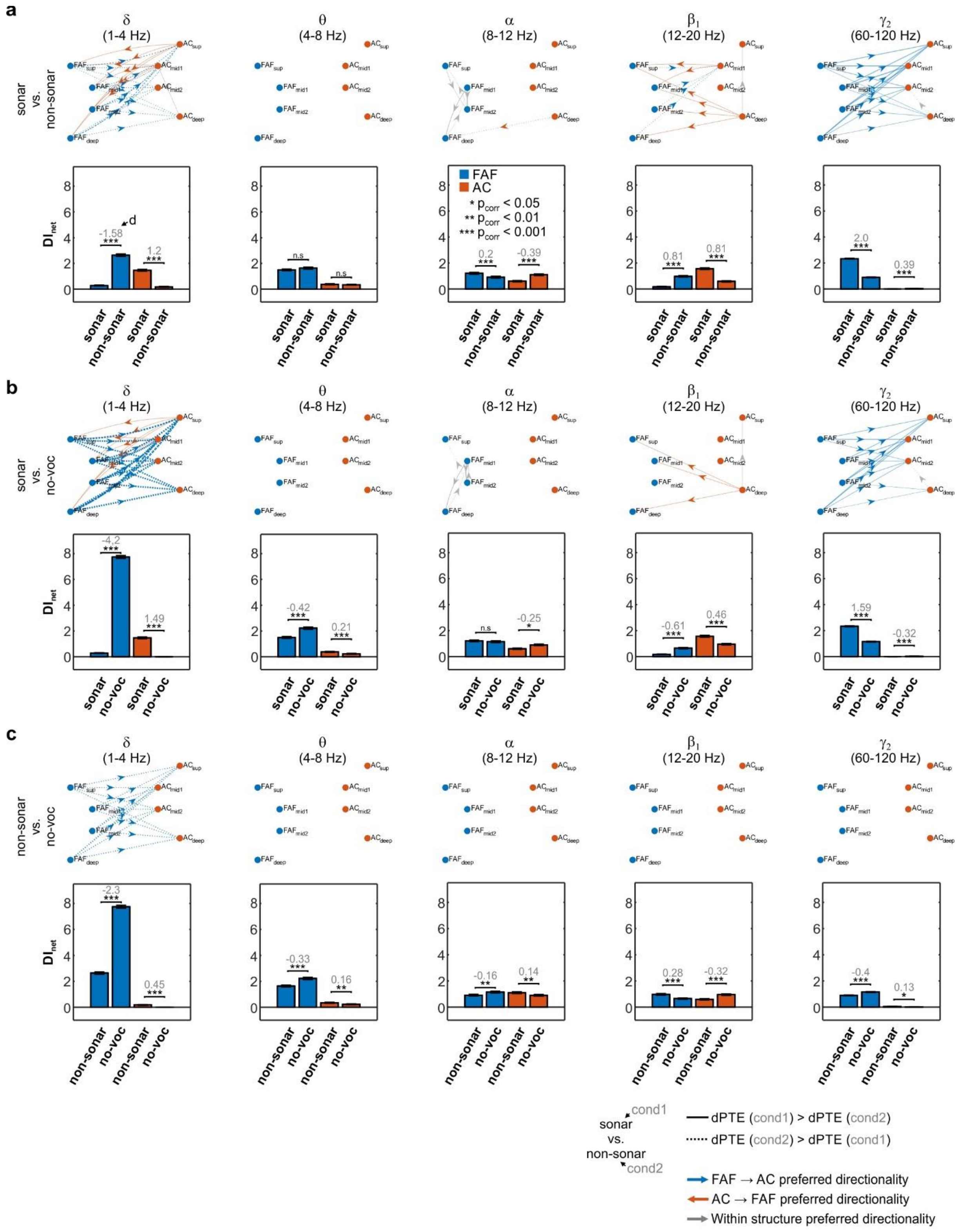
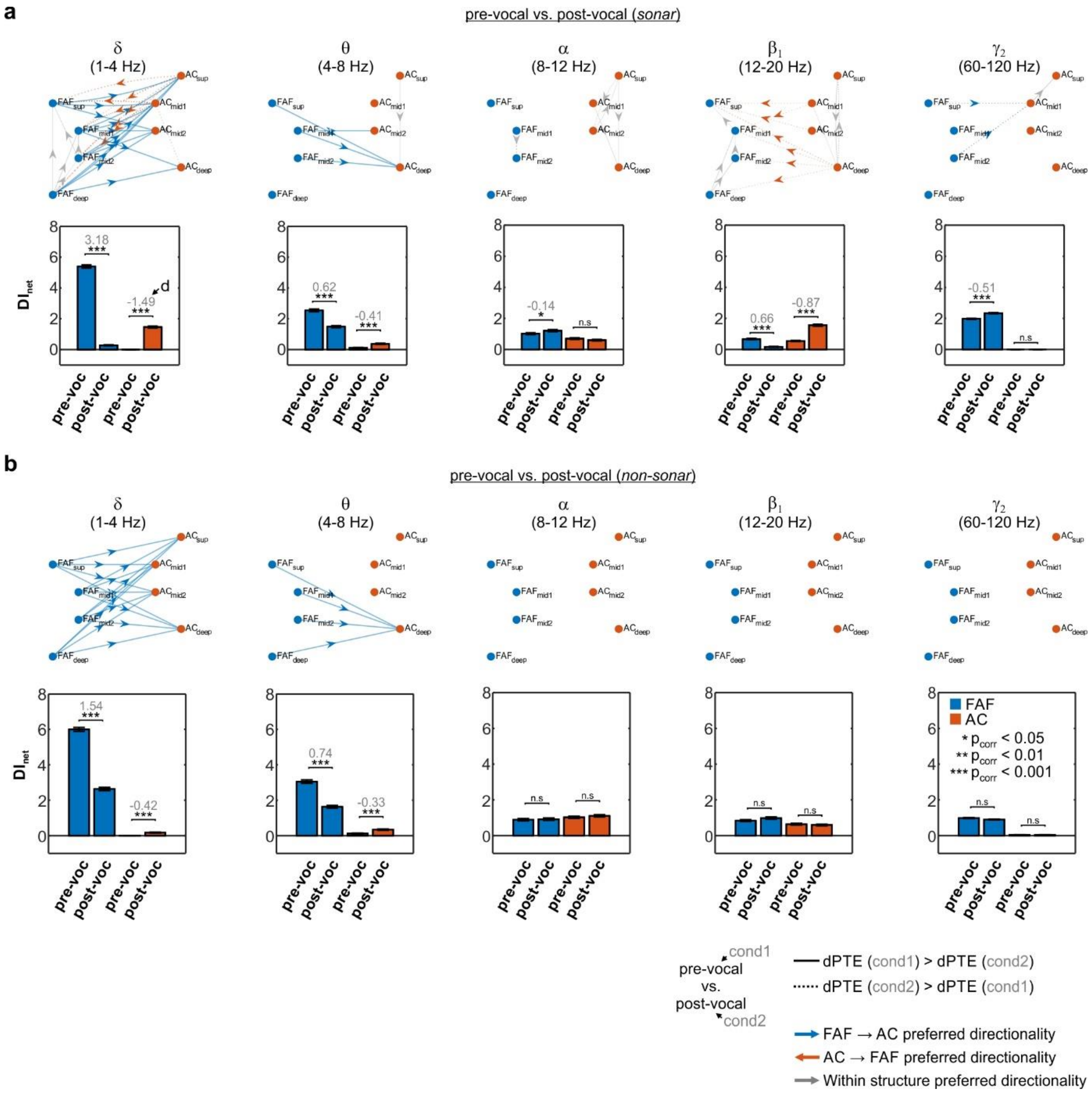


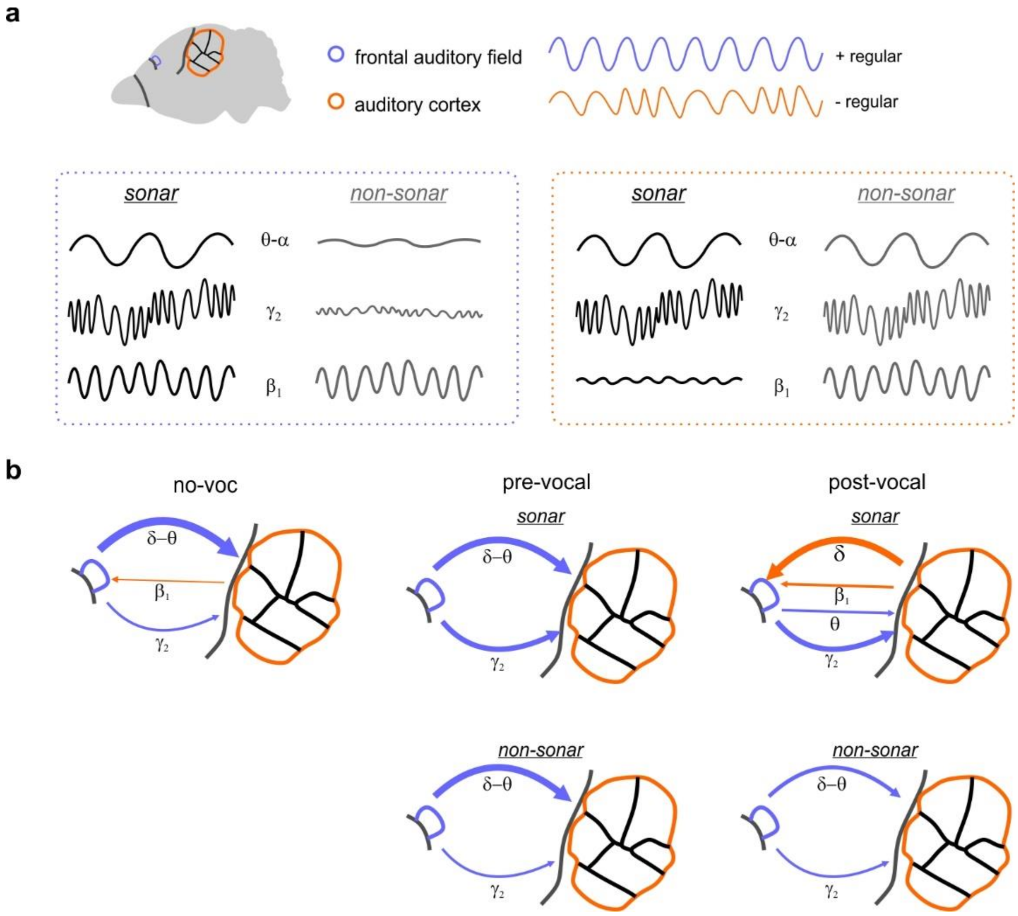
Fig. 5. Post-vocal dPTE differences across vocalization conditions.



1170

1171 **Fig. 6. Pre-vocal and post-vocal directionality differences in the FAF-AC network.**

1172



1174 **Fig. 7. The FAF-AC network during vocal production.**