Distinct cross-frequency coupling signatures for top-down and bottom-up information processing in auditory cortex

Christian D. Márton[1,2*¤], Makoto Fukushima[2,3,4], Corrie R. Camalier[2], Simon R. Schultz[1], Bruno B. Averbeck[2]

1 Centre for Neurotechnology & Department of Bioengineering, Imperial College London, London, SW7 2AZ, U.K.
2 Laboratory of Neuropsychology, Section on Learning and Decision Making, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA
3 Center for Brain Science, RIKEN, Saitama, Japan
4 Consumer Neuroscience, The Nielsen Company, Tokyo, Japan

¤Current Address: Laboratory of Neuropsychology, Section on Learning and Decision Making, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA
* cm3614@ic.ac.uk

Abstract

Specific frequency bands have been proposed to be predominantly employed in bottom-up and top-down information transfer (e.g. "γ up, β down"), but it remains unclear how and to what extent this notion could be generalized to cross-frequency interactions. We obtained auditory evoked field potentials recorded simultaneously from four auditory sectors of the macaque, and examined directed cross-frequency information transfer across the auditory hierarchy. Our findings showed that low frequencies (e.g. β) can be predominant bottom-up and high frequencies (e.g. γ) top-down, suggesting the assumption that sensory information is propagated mainly using distinct frequency ranges upwards and downwards need not always hold true. Rather, our findings revealed distinct signature patterns for bottom-up and top-down information processing among *cross*-frequency interactions. These patterns are largely preserved across coupling types (phase-amplitude and amplitude-amplitude) and across stimulus types (natural and synthetic auditory stimuli), suggesting they are a general hallmark of information processing in auditory cortex.

1 Introduction

One fundamental yet poorly understood component of cortical computation is the presence of widespread reciprocal cortical connections [Scott et al., 2015, Romanski and Averbeck, 2009, Kveraga et al., 2007, Salin and Bullier, 1995]. It remains unclear whether there are fundamental differences between bottom-up and top-down information processing. Here, we present a novel method to study directional interactions across the frequency spectrum, and employ it to contrast top-down and bottom-up cross-frequency interactions.

One prevalent theory of information processing in the brain which involves bottom-up and top-down components is *predictive coding*. In this (Bayesian) view of the brain, expectations are formed about incoming sensory information. These top-down predictions (expectations) are compared with bottom-up sensory information (outcomes) and representations are then potentially updated based on a prediction error (surprise), which is the difference between sensory inputs and their expectations. Neural implementations of predictive coding make specific predictions about top-down and bottom-up information transfer. Lower-order areas are thought to feed forward the

prediction error. At the level of high-order areas, this error - together with current predictions - is used to update expectations. The linear accumulation of these prediction errors is thought to result in the attenuation of higher frequencies, while the non-linear transformation of predictions will favor higher frequencies [Bastos et al., 2012, Friston, 2008]. This generates the prediction that lower frequencies (e.g. β) should be favored over higher frequencies (e.g. γ) in top-down information transfer, and vice versa in bottom-up information transfer. Several studies have supported this view [Michalareas et al., 2016, Bastos et al., 2015, Bastos et al., 2012, Fontolan et al., 2014], though not without exceptions.

Cross-frequency coupling is hypothesized to allow for top-down and bottom-up modulation of information processing in the brain. Cross-frequency coupling refers to the phase or amplitude component of a Fourier-transformed signal that is correlated with and possibly influences other components across the frequency spectrum. It may occur in several forms [Jirsa and Mueller, 2013], and has been observed during various tasks across several species including humans, macaques, and rodents. Coupling of low-frequency (e.g. θ, α) phase to high-frequency (e.g. γ) amplitude, in particular, has been associated with various behavioral mechanisms including learning, memory and attention [Hyafil et al., 2015b, Colgin, 2015, Doesburg et al., 2012, Kendrick et al., 2011, Canolty and Knight, 2010, Tort et al., 2009, Tort et al., 2008, Cohen, 2008, Jensen and Colgin, 2007, Canolty et al., 2006, Schack et al., 2002]. The phase of low-frequency oscillations could modulate the amplitude of high-frequency oscillations, putatively allowing control of spiking activity [Buzsáki et al., 2012, Ray and Maunsell, 2011]. Based on the predictive coding framework [Bastos et al., 2012, Friston, 2008], we would expect to observe the phase of the modulating (source) rhythm to be *lower* in frequency (e.g. θ,α,β) in the top-down direction, and *higher* in frequency (e.g. γ) in the bottom-up direction.

Although cross-frequency coupling can offer an important window unto information processing in the brain, current methods do not allow for directional components to be computed in a straightforward and computationally efficient manner. Previous studies and measures of cross-frequency coupling [Bruns and Eckhorn, 2004, Canolty et al., 2006, Cohen, 2008, Tort et al., 2010, Voytek et al., 2013, van Wijk et al., 2015] do not include direction in the estimation of coupling strength, thus making it impossible to distinguish feedforward and feedback components. Studies that do include directional measures [Michalareas et al., 2016, Bastos et al., 2015, Bastos et al., 2012, Fontolan et al., 2014] employ separate measures to look at coupling strength (i.e. through coherence) and directional information flow (i.e. through Granger-causality), and do not take *cross*-frequency interactions into account. The dynamic causal modeling (DCM) framework [Friston et al., 2017, Friston et al., 2003] does allow for the computation of directed coupling across frequencies [Furl et al., 2014], but is computationally expensive and requires a separate dimensionality reduction step before estimation.

Our approach provides an unified framework for estimating directional coupling strength. This framework allows for all cross-frequency interactions to be estimated simultaneously with an inherent dimensionality reduction. Thus, our method yields itself to the study of directed cross-frequency coupling in large-scale simultaneous recordings. We employed this method to systematically investigate the role of cross-frequency in bottom-up and top-down information processing. We obtained estimates of both phase-amplitude coupling (PAC) and amplitude-amplitude coupling (AAC) strength as measures of information processing in a particular direction.

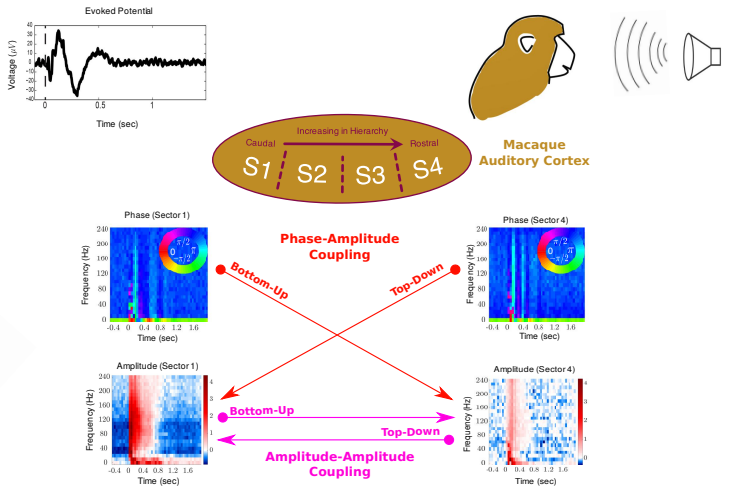
In the present study, we analyzed auditory evoked potentials recorded simultaneously from multiple auditory cortical areas of the macaque. These recordings span four sectors in auditory cortex, from *earlier*, caudally-located sectors to *later*, rostrally-oriented sectors [Scott et al., 2015, Hackett et al., 2014, Camalier et al.,

2012, Fukushima et al., 2012, Kikuchi et al., 2010]. We operationally defined feedforward to occur from earlier to later sectors (bottom-up direction), and feedback to occur from later to earlier sectors (top-down direction).

Our previous study showed that early caudal areas of primate auditory cortex retain high-fidelity responses for both natural vocalizations and synthetic sounds, but that later rostral areas preferentially retain information for natural vocalization only [Fukushima et al., 2014]. It is reasonable to infer that these coding differences would translate to systematic differences in coupling patterns between natural and synthetic stimuli. Alternatively, if the revealed coupling pattern is a more general hallmark of inter-areal communication, it should not be specific to stimulus type: natural and synthetic stimuli should show the same pattern of the couplings.

For both PAC and AAC, we found that top-down coupling was dominant when the frequency of the source sector was lower than that of the target sector. In contrast, bottom-up coupling was dominant when the source sector frequency was lower than that of the target sector. Specifically, PACs showed dominant *top-down* coupling of α phase in the source sector to broadband amplitude in the target sector, and dominant *bottom-up* coupling of multiple frequencies in the source sector to α amplitude in the target sector. AACs reflected the same pattern, but additionally included significant coupling asymmetries in the γ frequency range. These patterns were conserved across natural and synthetic stimuli, although overall reduced coupling strength was found for synthetic stimuli. Spectrally intact (but envelope-flattened) synthetic stimuli showed smaller differences in coupling strength to natural vocalization stimuli than did spectrally flat (but envelope preserved) synthetic stimuli. AACs showed more widespread differences in coupling strength than PACs among synthetic stimuli, in particular in the γ frequency range.

A



B

How strong is the coupling across regions, in the bottom-up and top-down direction?

Computing directional coupling strength

$$X = u_1 s_1 v_1^T, \quad \tilde{Y} = u_2 s_2 v_2^T$$

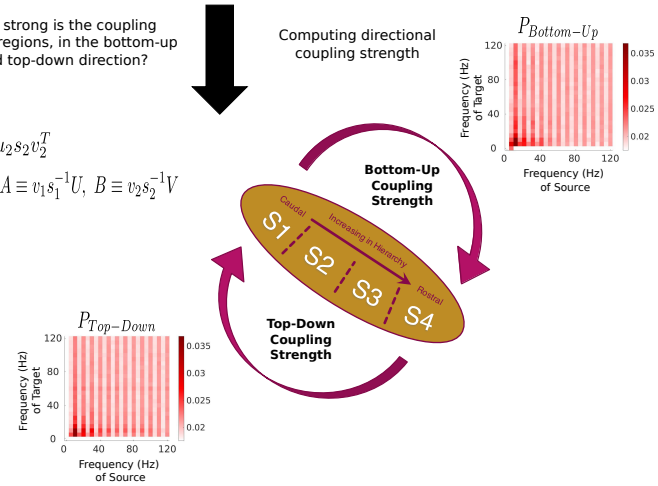
$$u_1^T u_2 = USV^T, \quad with \quad A \equiv v_1 s_1^{-1} U, \; B \equiv v_2 s_2^{-1} V$$

$$P = BSA^T$$

X : *Source Matrix*

$\tilde{Y}$: *Target Matrix*

P : *Coupling Matrix*



C

How much more control is exerted in one direction over the other?

Computing the difference between top-down and bottom-up coupling strength

$$\Delta Coupling\ Strength = \frac{1}{n}\sum_{i=1}^{n} \left| P^i_{Top-Down} \right| - \left| P^i_{Bottom-Up} \right|$$

red : $top-down \quad dominant$

$blue$: $bottom-up \quad dominant$

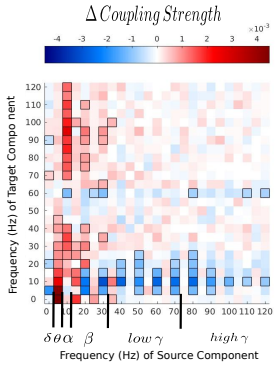


Fig 1. Coupling Analysis Overview: (A) Recordings were made from four μECoG arrays spanning auditory cortex while monkeys listened to one of 20 natural vocalizations (VOC), 20 synthetic envelope-preserved sounds (EPSs) or 20 synthetic spectrum-preserved sounds (SPSs). Electrodes were partitioned into four sectors along the caudorostral axis (S1: A1/ML (primary auditory cortex/middle lateral belt), S2: R (rostral core region of the auditory cortex / AL (anterior lateral belt region of the auditory cortex), S3: RTL (lateral rostrotemporal belt region of the auditory cortex), S4: RTp (rostrotemporal pole area)). We decomposed the signal into its time-frequency representation, and obtained amplitude and phase components for each sector. We investigated two types of coupling across sectors: amplitude-amplitude and phase-amplitude coupling. **(B)** We computed Granger-causal phase-amplitude and amplitude-amplitude coupling strength across the frequency spectrum within the framework of canonical correlation analysis. For each target, we regressed out the target's past together with the current value of the source and obtained the residual ($\widetilde{Y}$). We then computed cross-regional coupling strength for every source-target pair using the signal from the source sector together with the residual from the target sector. This was done separately in the bottom-up and in the top-down direction. Canonical correlation can be understood as performing a singular value decomposition (SVD) on the correlation matrix ($u_1^T u_2$), yielding canonical variables A and B. Couplings (P) are obtained by multiplying the singular values (S) of the correlation matrix with the canonical variables A and B (see Methods and Materials for details). **(C)** Bottom-up and top-down coupling strength was contrasted by computing the *difference* between top-down and bottom-up coupling matrices. The result was averaged across all source-target pairs (see Methods and Materials).

2 Methods and Materials

The dataset employed in the current analysis is the same as previously reported in Fukushima (2014) [Fukushima et al., 2014].

2.1 Subjects

Three adult male rhesus monkeys (Macaca mulatta) weighing 5.5–10 kg were used for recordings. All procedures and animal care were conducted in accordance with the Institute of Laboratory Animal Resources Guide for the Care and Use of Laboratory Animals. All experimental procedures were approved by the National Institute of Mental Health Animal Care and Use Committee.

2.2 Recordings

Custom-designed μECoG arrays (NeuroNexus Technologies, MI, USA) were used to record field potentials from macaque auditory cortex. Arrays were machine fabricated on a very thin polyimide film (20 μm), with each array featuring 32 recording sites, 50 μm in diameter each, on a 4 x 8 grid with 1 mm spacing (i.e., 3 x 7 mm rectangular grid) [Fukushima et al., 2014, Fukushima et al., 2012]. Two animals, monkeys B and K, were implanted with arrays each in the left hemisphere, while one animal, monkey M, was implanted with arrays in the right hemisphere. Three of the arrays in each monkey were placed on top of STP, in a caudorostrally oriented row [Fukushima et al., 2014]. The implantation procedure was described in detail previously [Fukushima et al., 2014, Fukushima et al., 2012].

2.3 Stimuli

The stimuli used for the main experiment included 20 conspecific monkey vocalizations (VOC) and two sets of 20 synthetic stimuli each (envelope-preserved sound, EPS; and spectrum preserved sound, SPS) derived from the original VOC stimuli. The VOC stimulus set consisted of 20 macaque vocalizations employed in a previous study [Kikuchi et al., 2010].

In order to obtain EPS from VOC stimuli, the envelope of a particular vocalization was estimated based on the amplitude component of the original stimulus' Hilbert transform. The amplitude envelope was then multiplied by broadband white noise to create the EPS stimulus. Thus, all 20 EPS stimuli exhibited flat spectral content; these stimuli could not be discriminated based on spectral features, while the temporal envelopes (and thus the durations) of the original vocalizations were preserved.

SPS stimuli were obtained by first generating broadband white noise with a duration of 500 ms and computing its Fourier transform. The SPS stimulus' amplitude in the Fourier domain was then replaced by the average amplitude of the corresponding VOC stimulus, before transforming back to the time domain by computing the inverse-Fourier transform. This resulted in a sound waveform that preserved the average spectrum of the original vocalization, while exhibiting a flat temporal envelope, random phase, with a duration of 500 ms. A 2 ms cosine rise/fall was then imposed on the stimulus to avoid abrupt onset/offset-effects. Hence, all 20 SPS stimuli exhibited nearly identical, flat temporal envelopes; these stimuli could not be discriminated using temporal features, while the average spectral power of the original vocalizations was preserved.

A total of 60 stimuli were presented in pseudorandom order, with an interstimulus interval of 3 s. Each stimulus was presented 60 times. The sound pressure levels of the stimuli measured by a sound level meter ranged from 65 to 72 dB at a location close to the animal's ear. Stimulus duration ranged from around .15 to 1 second [Fukushima et al., 2014].

2.4 Stimulus presentation and recording parameters

The experimental procedure was described in detail previously [Fukushima et al., 2014].The monkey was placed in a sound-attenuating booth for the experiment (Biocoustics Instruments). The sound stimuli were presented while the monkey sat in a primate chair and listened passively with its head fixed. Auditory evoked potentials from the 128 channels of the ECoG-array were bandpassed between 2-500 Hz, digitally sampled at 1500 Hz, and stored on hard-disk drives by a PZ2–128 preamplifier and the RZ2 base station (Tucker Davis Technology) .

2.5 Data Analysis

Data analysis was performed using MATLAB R2017a (MathWorks) software. Since there was little significant auditory evoked power above 250 Hz, recordings were low-pass filtered and resampled at 500 Hz to enhance calculation speed and reduce memory requirements. Data from each site was re-referenced by subtracting the average of all sites within the same array [Kellis et al., 2010].

The 96 sites on the STP were grouped based on the characteristic frequency maps obtained from the high-gamma power of the evoked response to a set of pure-tone stimuli [Fukushima et al., 2014,Fukushima et al., 2012]. This resulted in a grouping of electrodes into four sectors (Fig 1), which were estimated to correspond to the following subdivisions at a caudorostral-level within macaque auditory cortex: Sec (Sector) 1, A1/ML (primary auditory cortex/ middle lateral belt); Sec 2, R (rostral core region of the auditory cortex)/AL (anterior lateral belt region of the auditory cortex); Sec 3, RTL

(lateral rostrotemporal belt region of the auditory cortex); Sec 4, RTp (rostro-temporal pole area).

We employed the framework of canonical correlation analysis [Durstewitz, 2017] to compute Granger-causal interactions between frequency pairs. Thus, in order to assess gains in predictability from X to Y, beyond what is already known from Y_{past} and not just mere correlation, we started by regressing out Y's own past using $p = 10$ predictors (Eq 1). In addition, we also regressed out the current value of X_t; this was done to ensure that the result did not reflect instantaneous correlations between X and Y, but really a temporally predictive relationship. Subsequent analysis was performed using the residuals $\widetilde{Y}$ of the target region together with the source signal X.

$$\widehat{Y}_t = a_0 + \sum_{i=1}^{p} A_i Y_{t-i} + B_0 X_t, \quad \widetilde{Y}_t = Y_t - \widehat{Y}_t \tag{1}$$

In a second step, the field potential data was transformed into frequency space by computing the fast Fourier transform (FFT) of X and $\widetilde{Y}$ in steps of 200 ms, with a 100 ms overlap, using a rectangular windowing function, w. Subsequent analysis was carried out on the first 25 frequency bins of this data, ranging from 1 to 125 Hz in frequency space, which is well below the Nyquist-criterium.

We defined the following terminology to refer to particular frequency spaces throughout this paper: *delta* range: 0 - 2.5 Hz; *theta* range: 2.5 - 7.5 Hz; *alpha* range: 7.5 - 12.5 Hz; *low beta* range: 12.5 - 22.5 Hz; *high beta* range: 22.5 - 32.5 Hz; *low gamma* range: 32.5 - 72.5 Hz; *high gamma* range: $>$ 72.5 Hz. The delta frequency band (0 - 2.5 Hz) is not included in the analysis since the pre-amplifier employed during the recording sessions does not allow for recording frequency components below 2 Hz and the power in this band was expectedly low.

In order to compute phase-amplitude and amplitude-amplitude cross-frequency couplings across the four sectors (Fig 1A), the phase or amplitude component, respectively, of the Fourier transformed data of each recording electrode in one sector was regressed on the amplitude component of each recording electrode in another sector. For PAC, the phase component was sine and cosine transformed, and both were entered into the analysis. Coupling was computed using canonical correlation analysis, which transforms the data in such a way as to maximize the correlation between the inputs.

In canoncial correlation analysis, one first obtains the singular value decomposition (SVD) of input matrices $X = (x_{i,j})_{i=1,\ldots,n,j=1,\ldots,m}$ and $\widetilde{Y} = (\tilde{y}_{i,j})_{i=1,\ldots,n,j=1,\ldots,l}$ with n : number of observations, m : number of predictor variables, and l: number of regressor variables (Eq 2). X contained both the sine and cosine transformed phase component of the fast Fourier-transformed signal of a particular electrode in the source sector together with a columns of ones, while $\widetilde{Y}$ contained the squared and log-transformed amplitude component of the fast Fourier-transformed signal of an electrode in the target sector, with $n = 27600$ (1200 trials * 26 windows) for VOC and SPS stimuli and $n = 27531$ (1197 trials * 26 windows) for EPS stimuli, $l = 25$, and $m = 25 * 2$ in the case of PAC, whereas $m = 25$ in the case of AACs.

$$X = u_1 s_1 v_1^T, \quad \widetilde{Y} = u_2 s_2 v_2^T \tag{2}$$

Then, one obtains the canonical variables $A \equiv (a_{d1,d})_{i=1,\ldots,m,j=1,\ldots,k}$ and $B \equiv (b_{d2,d})_{i=1,\ldots,l,j=1,\ldots,k}$, with $k = min(rank(X), rank(\widetilde{Y}))$, by performing SVD on the correlation matrix $u_1^T u_2$ (in the space spanned by the observations contained in X and $\widetilde{Y}$) and projecting the resulting left and right unitary matrices into the phase and amplitude information space of X and $\widetilde{Y}$, respectively (Eq 3). Equivalently, one can obtain A and B by first performing a QR-decomposition of X and $\widetilde{Y}$, and performing

SVD on the correlation matrix of the q-components. The latter method was employed here.

$$u_1^T u_2 = USV^T, \quad with \quad A \equiv v_1 s_1^{-1} U,\ B \equiv v_2 s_2^{-1} V \tag{3}$$

One obtains the coupling strength matrix, P, by multiplying the B and A matrices with the singular values of the correlation matrix contained in S (Eq 4). For PAC, one combines the sine and cosine components of P by computing their Euclidian norm, yielding the final coupling matrix P^*_{PA}. For AAC, P is the final coupling matrix.

$$P = BSA^T, \quad P^*_{PA} = \begin{pmatrix} \sqrt{P^2_{d2,1} + P^2_{d2,m/2+1}} \\ ... \\ \sqrt{P^2_{d2,m/2} + P^2_{d2,m}} \end{pmatrix} \tag{4}$$

Directed coupling strength (Fig 1B) can thus be obtained by entering all of the time-phase (PAC) or time-amplitude data (AAC) as regressors (X) and time-amplitude data as targets (Y) simultaneously into just one computation that yields measures of coupling strength across all frequency pairs. We thus obtained estimates of coupling strength in two directions, bottom-up and top-down: in the former case, the source component was derived from a recording site in a sector lower in hierarchy than the target component (referred to as bottom-up coupling), while in the latter case, the source component came from a sector higher in hierarchy than the target component (referred to as top-down coupling). This analysis was performed for all channel pairs across all possible cross-regional pairs in each of the three monkeys (Fig 1B).

In order to obtain the difference ΔP between top-down and bottom-up coupling (Fig 1C), the magnitude (absolute value) of bottom-up coupling strength was subtracted from the magnitude of top-down coupling strength for each cross-regional pair (Eq 5). This was done for intact vocalizations (VOC) and synthetic stimuli (EPS, SPS) alike. Since the difference in coupling strength was similar across cross-regional pairs, results are presented averaged across all pairs ($n = 6$).

$$\Delta P = \frac{1}{n} \sum_{i=1}^{n} \mid P^i_{Top-Down} \mid - \mid P^i_{Bottom-Up} \mid \tag{5}$$

In order to compute coupling strength across time, a FFT with a window size of 100 ms and an overlap of 50 ms was used to transform the data into frequency space. Coupling strength was computed across 4 windows in order to achieve numerical stability, yielding a total of 11 windows of 200 ms length each. This resulted in input matrices X and Y of dimensionality $n = 1200 * 4$ trials for VOC and SPS stimuli and $n = 1197 * 4$ for EPS stimuli, $l = 13$, and $m = 13 * 2$ in the case of PACs, whereas $m = 13$ in the case AACs.

2.6 Statistical Analysis

Paired t-tests were performed to establish statistical significance of ΔP values, and of differences in coupling strength between intact and synthetic stimuli. These were performed separately for each frequency pair, including electrodes from across all cross-regional pairs and all three animals. In the case of couplings over time, paired t-tests were conducted separately for each frequency pair and time window, again including electrodes from across all cross-regional pairs and all three animals.

Bonferroni-correction was applied by dividing the significance threshold by the product of the number of source and target frequency bands employed ($p = 0.05/25^2$), and by the product of the number of target bands and time windows

($p = 0.05/(13 * 11)$), respectively. In addition, selected ΔP values were checked against surrogate values, which were constructed by scrambling the indices of top-down and bottom-up pairs and recomputing ΔP for 1000 repeats.

3 Results

We analyzed local field potentials recorded simultaneously from multiple cortical areas in the auditory cortex of three macaques while the animals listened to auditory stimuli (Fig 1A). These auditory stimuli consisted of 20 natural conspecific vocalizations and 40 synthetic sounds of two types, envelope-preserved stimuli (EPS) and spectrum-preserved stimuli (SPS) [Fukushima et al., 2014]. Our previous study established the approximate rostrocaudal location of implanted micro-electrocorticography (μECoG) arrays along the auditory hierarchy by the characteristic frequency of each electrode contact [Fukushima et al., 2014, Fukushima et al., 2012]. This partitioned the recording sites into four sectors (S1-S4), putatively spanning caudorostral level of core A1 (S1) to level of RTp (S4), including some of the surrounding belt areas (Fig1A).

First, the signal from every trial and every site was decomposed into its spectral components using a short-time Fourier transform. This resulted in time-phase and time-amplitude data (Fig 1A). Next, we obtained directed measures of coupling strength by computing Granger-causal coupling coefficients within a canonical correlation framework (see Methods and Materials). We computed directional coupling strength across the frequency spectrum for each cross-regional source-target pair in the top-down and bottom-up direction (Fig 1B). Top-down coupling was defined as coupling in which the modulating source (i.e. phase band) originated in a sector of higher order than the modulated target (i.e. amplitude band): for example, coupling between the phase in S4 and the amplitude in S1 is defined as top-down, and coupling between the phase in S1 and the amplitude in S4 is defined as bottom-up. We included all cross-regional pairs across the four sectors in the analysis.

Finally, In order to assess how much stronger coupling was in one direction over the other, we computed the difference between top-down and bottom-up coupling strength (Fig 1C). The analyses are presented collapsed across all cross-regional pairs as we focused on examining signatures of top-down and bottom-up coupling that are shared across the auditory hierarchy.

3.1 Distinct cross-frequency coupling signatures for top-down and bottom-up processing

We first examined directional PAC and AAC across the four sectors in auditory evoked potentials to 20 natural conspecific vocalizations (VOC). Both phase-amplitude and AACs peaked in the α frequency range (Fig 2A,B,D,E). The field potential signal showed highest power in the α frequency range (Fig S1), decaying thereafter with the typical $1/f$ pattern. The modulating phase band (α frequency range, 5-10Hz) thus includes the peak in power spectrum [Aru et al., 2015]. PAC shows a banding pattern originating from the harmonics of the fundamental phase (Fig 2A, B), whereas AAC does not show any such banding pattern (Fig 2D, E).

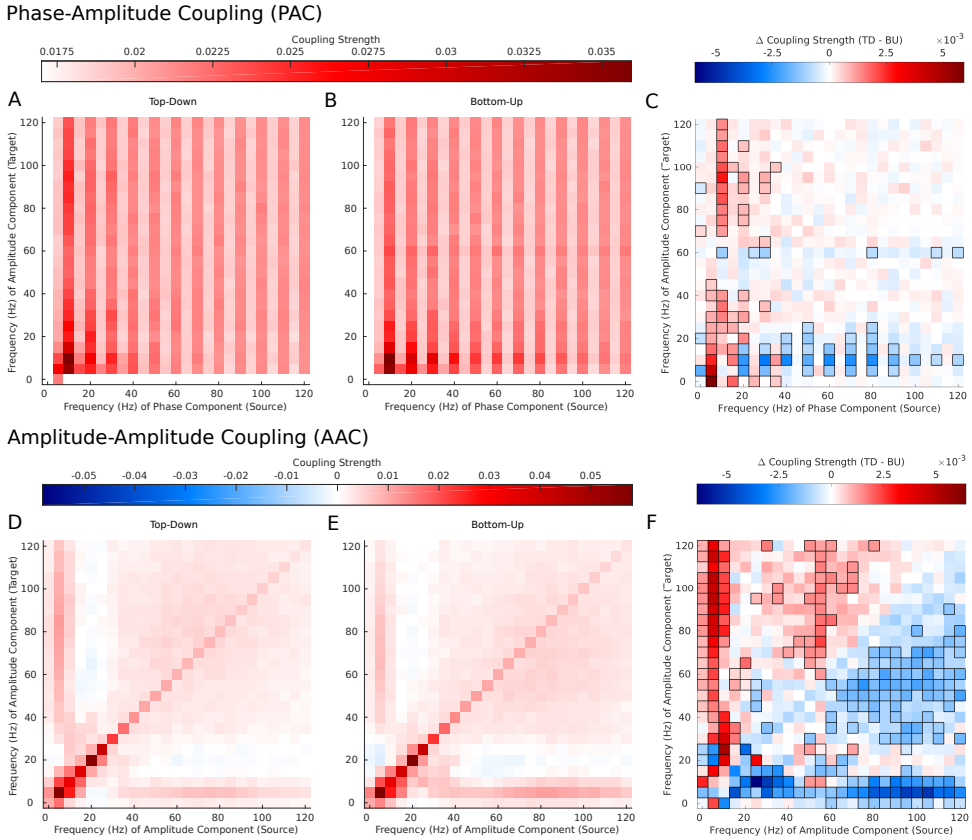


Fig 2. Top-down vs. bottom-up phase-amplitude and amplitude-amplitude coupling in natural vocalizations (VOC). (A-B) Phase-amplitude coupling (PAC) strength in the top-down (A) and bottom-up direction (B). Coupling strength decreases with increasing frequency, across top-down and bottom-up pairs. Depicted are canonical correlation-derived coupling coefficients, averaged across all channels, animals and cross-regional pairs (see Methods and Materials). **(C)** Difference in top-down and bottom-up in PAC strength. Significant differences ($p < 1.9223e - 5$) are enclosed by black rectangles. **(D-E)** Amplitude-amplitude coupling (AAC) strength in the top-down (D) and bottom-up direction (E). Depicted are canonical correlation-derived coupling coefficients, averaged across all channels, animals and cross-regional pairs. **(F)** Difference between top-down and bottom-up in AAC strength. Significant differences ($p < 1.9223e - 5$) are enclosed by black rectangles.

Next, we examined differences between top-down and bottom-up coupling for PAC and AAC (Fig 2C, G). For both types of coupling, we found that many off-diagonal components (*cross*-frequency couplings) showed a significant difference between the two directions, as indicated by black rectangles in the figures (p<1.9223e-5). Overall, top-down coupling strength was dominant in the upper diagonal region (as indicated by red colors) whereas bottom-up was dominant in the lower diagonal region (as indicated by blue colors). In other words, top-down coupling dominated when the frequency of the source sector was lower than that of the target sector, while bottom-up coupling dominated when the frequency of the source sector was higher than that of the target sector.

Among PACs, top-down coupling strength was dominant for the combination of low source frequency and high target frequency (Fig 2C). In particular, top-down coupling was dominant between α phase and broadband amplitude. We confirmed the significance of the difference in coupling strength for selected frequency pairs in the α

phase range by computing surrogate values: real values were found to be well to the right of the distribution of surrogate values, confirming the significance of the observed effect. Additionally, top-down coupling was dominant between θ phase and α/β amplitude, α phase and β/γ amplitude, as well as between β phase and γ amplitude. The coupling of β phase to α amplitude was stronger top-down than bottom-up, and this was the only case in which the frequency of the source sector was higher than that of the target sector. In addition, top-down coupling was also dominant for selected same-frequency couplings (θ-θ and β-β PAC). Bottom-up coupling, in contrast, was dominant between high source frequency and low target frequency pairs (Fig 2C, F). More specifically, bottom-up coupling was dominant between β/γ phase and α/β amplitude. Dominant bottom-up coupling of β phase to α/β amplitude could be seen as the originator of coupling at higher harmonically related phases.

AACs showed a similar pattern to PACs in both directions, but additionally displayed γ range asymmetries. AACs replicated the pattern oveserved in PACs with dominant α to broadband coupling top-down, and dominant α to β coupling bottom-up (Fig 2F). Unlike PACs however, AACs also showed stronger top-down coupling of low γ band amplitude to high γ amplitude (Fig 2G), and stronger bottom-up coupling of high γ amplitude in the source sector to high γ amplitude in the target sector (Fig 2F).

3.2 Top-down and bottom-up cross-frequency signatures generalize to synthetic stimuli

To examine whether the observed coupling patterns are a more general hallmark of inter-areal communication, we also analyzed coupling strength in two types of synthetic stimuli derived from the original vocalizations. Envelope-preserved sounds (EPS) were obtained by leaving the original temporal envelope of the vocalization intact but flattening its spectral content, while spectrum-preserved sound (SPS) were obtained by leaving the spectral information of the original vocalization intact but flattening its temporal envelope [Fukushima et al., 2014].

Among PACs, EPS stimuli showed similar coupling patterns to intact vocalization stimuli (VOC), but displayed overall lower couping strength (Fig 3). As in VOC stimuli, we found lower frequency phase (θ,α) to higher frequency amplitude (β,γ) coupling to be stronger top-down, whereas low phase (β) to low amplitude (α) coupling was stronger bottom-up (Fig 3A). Evaluating the difference in coupling strength between VOC and EPS separately in the top-down and bottom-up direction, we found overall weaker coupling strength in EPS stimuli in both directions (Fig 3B,C). In the top-down direction, the largest difference between VOC and EPS stimuli occurred at α source phase (Fig 3B); at α phase, coupling strength for VOC was markedly enhanced for α, β, and γ amplitude couplings. In the bottom-up direction, there was little to no difference between the two stimuli in the β and γ amplitude range (Fig 3C); rather, the largest difference in coupling strength between VOC and EPS stimuli occurred at α target amplitude.

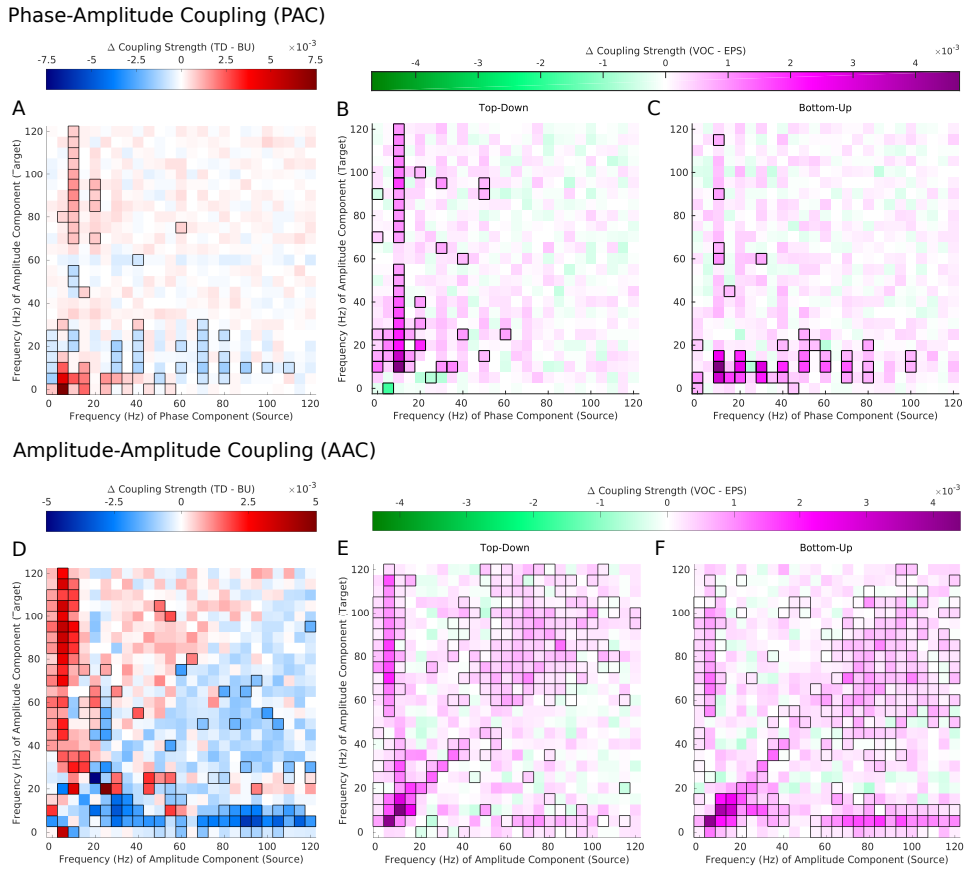


Fig 3. Top-down versus bottom-up phase-amplitude and amplitude-amplitude coupling strength in synthetic envelope-preserved sounds (EPS). (A) Difference in top-down and bottom-up *phase-amplitude* coupling (PAC) strength. Significant differences ($p < 1.9223e-5$) are enclosed by black rectangles. **(B-C)** Difference in PAC strength between natural (VOC) and synthetic envelope-preserved sounds (EPS), separately in the top-down (B) and bottom-up direction (C). Significant differences ($p < 1.9223e-5$) are enclosed by black rectangles. **(D)** Difference in top-down and bottom-up *amplitude-amplitude* coupling (AAC) strength. **(E-F)** Difference in AAC strength between VOC and EPS, separately in the top-down (E) and bottom-up direction (F).

Among AACs, EPS stimuli showed similar coupling patterns to intact stimuli (VOC) along the α amplitude axes but much reduced asymmetries in the γ frequency space (Fig 3D). Evaluating the difference in coupling strength between VOC and EPS separately in the top-down and bottom-up direction, AACs showed overall weaker coupling strength in EPS stimuli in both directions. Differences in coupling strength among AACs occurred across a broader frequency range than among PACs (Fig 3E,F). In particular, top-down and bottom-up coupling strength in the γ frequency range was weaker for EPS compared to VOC stimuli. This resulted in overall less top-down vs bottom-up asymmetries in the γ frequency range among EPS stimuli compared to VOC stimuli (Fig 3D).

SPS stimuli showed similar coupling patterns to VOC stimuli among PACs, and coupling strength was overall more similar between SPS and VOC than between EPS and VOC (Fig 4,A, D). In the top-down direction (Fig 4D), PAC strength was higher for VOC than for SPS stimuli among a few selected couplings in the high frequency range (γ), whereas PAC strength was higher for SPS than for VOC in the low frequency

region (among θ, α, β phase to α amplitude couplings; Fig 4B). In the bottom-up direction, there was little to no difference in PAC strength between VOC and SPS stimuli, especially in the α amplitude range (Fig 4C) where there were differences between VOC and EPS stimuli (Fig 3C).

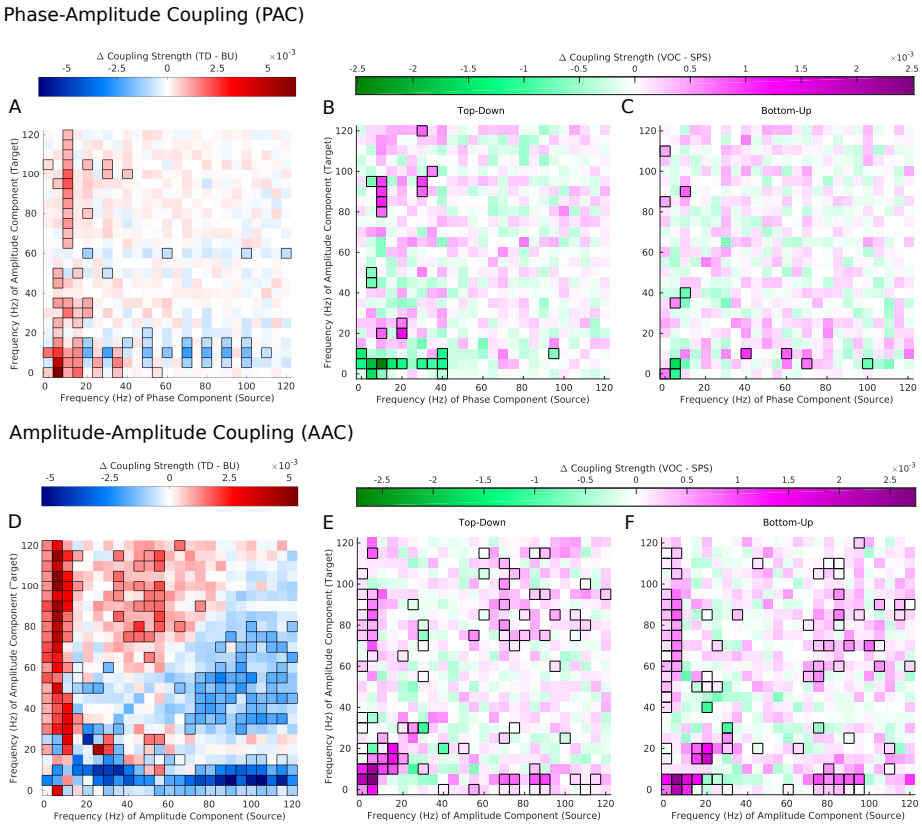


Fig 4. Top-down versus bottom-up phase-amplitude and amplitude-amplitude coupling strength in synthetic spectrum-preserved sounds (SPS). (A) Difference in top-down and bottom-up phase-amplitude coupling (PAC) strength. Significant differences ($p < 1.9223e-5$) are enclosed by black rectangles. **(B-C)** Difference in PAC strength between natural (VOC) and synthetic spectrum-preserved sounds (SPS), separately in the top-down (B) and bottom-up direction (C). Significant differences ($p < 1.9223e-5$) are enclosed by black rectangles. **(D)** Difference in top-down and bottom-up *amplitude-amplitude* coupling (AAC) strength. **(E-F)** Difference in AAC strength between VOC and SPS, separately in the top-down (E) and bottom-up direction (F).

The same was true among AACs, though there were more widespread differences in coupling strength than among PACs. There were more differences in coupling strength between SPS and VOC stimuli among AACs in the low frequency (θ, α, and β) and in the high frequency (γ) range (Fig 4D, F).

Overall, top-down and bottom-up cross-frequency coupling signatures were preserved among synthetic stimuli, albeit at overall reduced coupling strength. Envelope-flattened (but spectrally intact) SPS stimuli showed smaller differences in coupling strength to VOC stimuli than spectrally-flat (but envelope preserved) EPS stimuli. AACs showed more widespread differences in coupling strength than PACs among synthetic stimuli, in particular in the γ frequency range.

4 Discussion

We revealed distinct cross-frequency signatures for top-down and bottom-up information processing in auditory cortex. These signature patterns were largely preserved across phase-amplitude and amplitude-amplitude (though the latter additionally included significant γ range asymmetries), and across stimulus types (albeit at lower coupling strength in synthetic stimuli). This suggests the observed cross-frequency signatures could be a general hallmark of bottom-up and top-down information processing in the brain.

The top-down coupling profile we observed is in accordance with previous results, but we found more broad coupling in the bottom-up direction than previously reported. Previous studies generally stress that feedforward (or bottom-up) influences are mediated by gamma frequency range rhythms [Michalareas et al., 2016, Bastos et al., 2015, Fontolan et al., 2014, Bosman et al., 2012]. However, across each of these studies, we can find examples of a broader frequency range being employed for bottom-up information transmission. Fontolan et al. [Fontolan et al., 2014] report higher bottom-up coupling of delta (1-3Hz) frequency phase in left A1 to high gamma (80-100Hz) amplitude in left association auditory cortex using human depth recordings. Bastos et al. [Bastos et al., 2015] find that theta-channels are being used in addition to gamma-frequency channels in feedforward information transmission among primate visual areas. Finally, Michalareas et al. [Michalareas et al., 2016] do not observe stronger alpha-beta frequency range feedback (top-down) modulation in 8 out of 21 primate visual cortex area pairs (with 2 pairs showing the opposite pattern), and no stronger gamma frequency range feedforward (bottom-up) modulation in 5 out of 21 pairs. Similarly, we found α-α and α-*low* β amplitude-amplitude and phase-amplitude couplings to be equally strong in both directions. We also observed stronger bottom-up coupling in the low frequency space among both types of coupling (β-to-α/β). Among amplitude-amplitude couplings, we also found overall stronger top-down coupling in the γ frequency range. Differences in exact frequency ranges reported may be due to regional (visual versus auditory cortex) variation [Tort et al., 2009, Canolty et al., 2006] and species-related variation. Overall, these findings show there is no clear association of particular frequency bands with direction, pointing to a picture of top-down and bottom-up processing that is more complex than the previously hypothesized 'β-down vs. γ-up' [Bastos et al., 2012, Fontolan et al., 2014].

Rather, our findings point to the existence of generalized cross-frequency signatures of bottom-up and top-down information processing. This is expressed in dominant *top-down* coupling of α phase or amplitude in the source sector to broadband amplitude in the target sector, and dominant *bottom-up* coupling of multiple frequencies in the source sector to α amplitude in the target sector. By coupling to high-frequency rhythms, low frequency oscillations can shape neuronal excitability directly: high-frequency (e.g. γ range) oscillations have been shown to correlate with spiking events. It is also possible for low-frequency oscillations to shape neuronal excitability more indirectly, by coupling to low-frequency rhythms. Through a hierarchically organized oscillation structure, these low-frequency rhythms, in turn, can couple to high-frequency rhythms. Our findings suggest top-down processes involve control of spiking activity (through high frequency amplitude coupling), while bottom-up processes mainly exert influence more indirectly (through low frequency amplitude coupling).

We also observed the same pattern of results in synthetic stimuli. The overall lower coupling strength we observed in synthetic stimuli could be a consequence of the coding differences that were previously observed between synthetic and intact stimuli [Fukushima et al., 2014]. We observed more widespread differences in coupling strength across the frequency space between the two stimulus types among amplitude-amplitude than among phase-amplitude couplings. γ frequency range differences in coupling

strength among amplitude-amplitude couplings in particular appear to be more sensitive to acoustic features. Therefore, γ range amplitude-amplitude coupling could underlie encoding of stimulus features, whereas phase-amplitude couplings may reflect a more general hallmark of inter-areal communication.

The observation of overall decreased coupling strength in spectrally flat stimuli (EPS) suggests enhanced encoding and processing of information-rich, environmentally relevant stimuli such as the natural vocalization employed here. The fact that envelope-flattened (and spectrally rich) stimuli elicit a coupling profile in close resemblance to natural vocalization stimuli (VOC) suggests some aspects of auditory processing are relatively unhindered by envelope manipulations. This is also consistent with our previous results that showed information about spectrally rich stimuli (SPS) was better maintained than information about spectrally flat stimuli (EPS), at least across the first two sectors (S1,S2) of the auditory hierarchy. In the highest auditory sector (S4), decoding performance showed differences between VOC and EPS across a broader frequency range than between VOC and SPS [Fukushima et al., 2014].

Bottom-up PAC can be relevant to stimulus encoding. For example, an oscillation-based theory of speech encoding [Hyafil et al., 2015a, Giraud and Poeppel, 2012] predicts that cortical theta oscillations track the syllabic rhythm of speech and, in turn, reset the spiking of neurons at a gamma frequency level through theta-gamma PAC. This ensures that neural excitability is optimally aligned with incoming speech. Our finding of low frequency phase to high frequency amplitude coupling fits within a framework that requires theta-gamma coupling for accurate speech encoding [Hyafil et al., 2015a]. Our finding that these patterns are consistent across stimulus classes implies that this may be a generalized mechanism of auditory information processing which is co-opted during speech processing. Nevertheless, our finding of enhanced low frequency phase to high frequency amplitude coupling in the top-down direction shows these couplings also play a prominent role in top-down control - an aspect which should be incorporated into modeling studies.

Coupling in the top-down direction has been of particular interest in theories of the effect of attention on auditory information processing. Attention is thought to to sample stimuli rhythmically, fluctuating at a 4-10 Hz rhythm [Landau and Fries, 2012], and was shown to enhance the processing of degraded speech in the anterior and posterior STS [Wild et al., 2012]. Frontal top-down signals in the delta and theta frequency range have been shown to modulate low-frequency oscillations in human auditory cortex, increasing their coupling to continuous speech [Park et al., 2015]. In visual cortex, delta frequency oscillations have been shown to entrain with the rhythm of stimuli in the attended stream [Schroeder et al., 2010, Lakatos et al., 2008], and top-down beta-band modulation was found to be enhanced with attention [Bastos et al., 2015]. Increased power in low-frequency range oscillations may also, in turn, enhance bottom-up rhythms through top-down (cross-frequency) coupling [Bressler and Richter, 2015, Lee et al., 2013]; indeed, top-down beta-band activity was found to be maximally correlated with bottom-up gamma-band activity when top-down preceded bottom-up activity [Richter et al., 2017]. In addition, top-down couplings may also reflect the modulatory activity of predictive processes on speech perception & processing [Arnal and Giraud, 2012, Gagnepain et al., 2012, Davis et al., 2011, Wild et al., 2010, Poeppel et al., 2008]. Both attentive and predictive processes may be expressed through the prominent low frequency phase to high frequency amplitude coupling we observed here.

Comparisons of bottom-up and top-down coupling strength are relevant to models of neural information processing such as predictive coding. In this framework, prediction errors and predictions may thus be propagated up and down the cortical hierarchy, respectively, through cross-frequency couplings. Low frequency (θ,α,β) phase to high frequency amplitude (γ) couplings could be used mainly to propagate predictions

top-down, adjusting lower-order receptivity according to expectations. Meanwhile, prediction errors can influence predictions by interfering with low frequency (θ,α,β) oscillations in higher order areas through β-α interactions.

In summary, we revealed distinct cross-frequency signatures of top-down and bottom-up information processing. These signatures are largely preserved across coupling types and stimulus types, suggesting they are a general hallmark of information processing in auditory cortex. Our findings also call into question the notion that information is propagated mainly high frequencies bottom-up and low frequencies. They suggest the incorporation of cross-frequency effects into models of information processing in the brain such as predictive coding. We also employed a novel method that allows for the efficient computation of directed cross-frequency coupling strength in a single common framework. This method is particularly well-suited to the analysis of directional coupling strength across multiple areas in large-scale simultaneous recordings.

5 Acknowledgments

This work was supported by a Wellcome Trust NIH-fellowship (C.D.M. & S.S.) and by NIH grant ZIA MH002928-01 (B.B.A.).

6 Competing Interests

The authors report no conflicts of interest.

References

Arnal and Giraud, 2012. Arnal, L. H. and Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7):390–398.

Aru et al., 2015. Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., Singer, W., and Vicente, R. (2015). Untangling cross-frequency coupling in neuroscience. *Current Opinion in Neurobiology*, 31:51–61.

Bastos et al., 2012. Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, 76(4):695–711.

Bastos et al., 2015. Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., and Fries, P. (2015). Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron*, 85(2):390–401.

Bosman et al., 2012. Bosman, C. A., Schoffelen, J.-M., Brunet, N., Oostenveld, R., Bastos, A. M., Womelsdorf, T., Rubehn, B., Stieglitz, T., De Weerd, P., and Fries, P. (2012). Attentional Stimulus Selection through Selective Synchronization between Monkey Visual Areas. *Neuron*, 75(5):875–888.

Bressler and Richter, 2015. Bressler, S. L. and Richter, C. G. (2015). ScienceDirect Interareal oscillatory synchronization in top-down neocortical processing. *Current Opinion in Neurobiology*, 31:62–66.

Bruns and Eckhorn, 2004. Bruns, A. and Eckhorn, R. (2004). Task-related coupling from high- to low-frequency signals among visual cortical areas in human subdural recordings. *International Journal of Psychophysiology*, 51(2):97–116.

Buzsáki et al., 2012. Buzsáki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13(6):1–14.

Camalier et al., 2012. Camalier, C. R., DAngelo, W. R., Sterbing-DAngelo, S. J., de la Mothe, L. A., and Hackett, T. A. (2012). Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *PNAS*, 109(44):18168–18173.

Canolty et al., 2006. Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2006). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science*, 313(5793):1626–1628.

Canolty and Knight, 2010. Canolty, R. T. and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11):506–515.

Cohen, 2008. Cohen, M. X. (2008). Assessing transient cross-frequency coupling in EEG data. *Journal of Neuroscience Methods*, 168(2):494–499.

Colgin, 2015. Colgin, L. L. (2015). Theta–gamma coupling in the entorhinal–hippocampal system. *Current Opinion in Neurobiology*, 31:45–50.

Davis et al., 2011. Davis, M. H., Ford, M. A., Kherif, F., and Johnsrude, I. S. (2011). Does Semantic Context Benefit Speech Understanding through. *Journal of Cognitive Neuroscience*, pages 1–21.

Doesburg et al., 2012. Doesburg, S. M., Green, J. J., McDonald, J. J., and Ward, L. M. (2012). Theta modulation of inter-regional gamma synchronization during auditory attention control. *Brain Research*, 1431(C):77–85.

Durstewitz, 2017. Durstewitz, D. (2017). Advanced Data Analysis in Neuroscience. *Springer International Publishing*.

Fontolan et al., 2014. Fontolan, L., Morillon, B., Liegeois-Chauvel, C., and Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature Communications*, 5:1–10.

Friston, 2008. Friston, K. (2008). Hierarchical Models in the Brain. *PLOS Computational Biology*, 4(11):e1000211–24.

Friston et al., 2003. Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302.

Friston et al., 2017. Friston, K. J., Preller, K. H., Mathys, C., Cagnan, H., Heinzle, J., Razi, A., and Zeidman, P. (2017). Dynamic causal modelling revisited. *NeuroImage*, pages 1–0.

Fukushima et al., 2012. Fukushima, M., Saunders, R. C., Leopold, D. A., Mishkin, M., and Averbeck, B. B. (2012). Spontaneous high-gamma band activity reflects functional organization of auditory cortex in the awake macaque. *Neuron*, 74(5):899–910.

Fukushima et al., 2014. Fukushima, M., Saunders, R. C., Leopold, D. A., Mishkin, M., and Averbeck, B. B. (2014). Differential Coding of Conspecific Vocalizations in the Ventral Auditory Cortical Stream. *Journal of Neuroscience*, 34(13):4665–4676.

Furl et al., 2014. Furl, N., Coppola, R., Averbeck, B. B., and Weinberger, D. R. (2014). Cross-Frequency Power Coupling Between Hierarchically Organized Face-Selective Areas. *Cerebral cortex (New York, N.Y. : 1991)*, 24(9):2409–2420.

Gagnepain et al., 2012. Gagnepain, P., Henson, R. N., and Davis, M. H. (2012). Temporal Predictive Codes for Spoken Words in Auditory Cortex. *Current Biology*, 22(7):615–621.

Giraud and Poeppel, 2012. Giraud, A.-L. and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4):511–517.

Hackett et al., 2014. Hackett, T. A., de la Mothe, L. A., Camalier, C. R., Falchier, A., Lakatos, P., Kajikawa, Y., and Schroeder, C. E. (2014). Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. *Frontiers in Neuroscience*, 8(72):1–21.

Hyafil et al., 2015a. Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., and Giraud, A.-L. (2015a). Speech encoding by coupled cortical theta and gamma oscillations. *eLife*, 4:3958.

Hyafil et al., 2015b. Hyafil, A., Giraud, A.-L., Fontolan, L., and Gutkin, B. (2015b). Neural Cross-Frequency Coupling: Connecting Architectures, Mechanisms, and Functions. *Trends in Neurosciences*, 38(11):725–740.

Jensen and Colgin, 2007. Jensen, O. and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11(7):267–269.

Jirsa and Mueller, 2013. Jirsa, V. and Mueller, V. (2013). Cross-frequency coupling in real and virtual brain networks. *Frontiers in Computational Neuroscience*, pages 1–25.

Kellis et al., 2010. Kellis, S., Miller, K., Thomson, K., Brown, R., House, P., and Greger, B. (2010). Decoding spoken words using local field potentials recorded from the cortical surface. *Journal of Neural Engineering*, 7(5):056007–20.

Kendrick et al., 2011. Kendrick, K. M., Zhan, Y., Fischer, H., Nicol, A. U., Zhang, X., and Feng, J. (2011). Learning alters theta amplitude, theta-gamma coupling and neuronal synchronization in inferotemporal cortex. *BMC Neuroscience*, 12(1):55.

Kikuchi et al., 2010. Kikuchi, Y., Horwitz, B., and Mishkin, M. (2010). Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(39):13021–13030.

Kveraga et al., 2007. Kveraga, K., Ghuman, A. S., and Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, 65(2):145–168.

Lakatos et al., 2008. 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(5872):110–113.

Landau and Fries, 2012. Landau, A. N. and Fries, P. (2012). Attention Samples Stimuli Rhythmically. *Current Biology*, 22(11):1000–1004.

Lee et al., 2013. Lee, J. H., Whittington, M. A., and Kopell, N. J. (2013). Top-Down Beta Rhythms Support Selective Attention via Interlaminar Interaction: A Model. *PLOS Computational Biology*, 9(8):e1003164–23.

Michalareas et al., 2016. Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., and Fries, P. (2016). Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron*, 89(2):384–397.

Park et al., 2015. Park, H., Ince, R. A. 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. *CURBIO*, 25(12):1649–1653.

Poeppel et al., 2008. Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493):1071–1086.

Ray and Maunsell, 2011. Ray, S. and Maunsell, J. H. R. (2011). Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLoS Biology*, 9(4):e1000610–15.

Richter et al., 2017. Richter, C. G., Thompson, W. H., Bosman, C. A., and Fries, P. (2017). Top-Down Beta Enhances Bottom-Up Gamma. *Journal of Neuroscience*, 37(28):6698–6711.

Romanski and Averbeck, 2009. Romanski, L. M. and Averbeck, B. B. (2009). The Primate Cortical Auditory System and Neural Representation of Conspecific Vocalizations. *Annual Review of Neuroscience*, 32(1):315–346.

Salin and Bullier, 1995. Salin, P.-A. and Bullier, J. (1995). Corticocortical Connections in the Visual System: Structure and Function. *Physiological Review*, 75(1):1–48.

Schack et al., 2002. Schack, B., Vath, N., Petsche, H., Geissler, H. G., and Moeller, E. (2002). Phase-coupling of theta-gamma EEG rhythms during short-term memory processing. *International Journal of Psychophysiology*, 44:143–163.

Schroeder et al., 2010. Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, 20(2):172–176.

Scott et al., 2015. Scott, B. H., Leccese, P. A., Saleem, K. S., Kikuchi, Y., Mullarkey, M. P., Fukushima, M., Mishkin, M., and Saunders, R. C. (2015). Intrinsic Connections of the Core Auditory Cortical Regions and Rostral Supratemporal Plane in the Macaque Monkey. *Cerebral cortex (New York, N.Y. : 1991)*, 7:bhv277–32.

Tort et al., 2010. Tort, A. B. L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring Phase-Amplitude Coupling Between Neuronal Oscillations of Different Frequencies. *Journal of Neurophysiology*, 104(2):1195–1210.

Tort et al., 2009. Tort, A. B. L., Komorowski, R. W., Manns, J. R., Kopell, N. J., and Eichenbaum, H. (2009). Theta–gamma coupling increases during the learning of item– context associations. *PNAS*, 106(49):20942–20947.

Tort et al., 2008. Tort, A. B. L., Kramer, M. A., Thorn, C., Gibson, D. J., Kubota, Y., Graybriel, A. M., and Kopell, N. J. (2008). Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. *PNAS*, 105(51):20517–20522.

van Wijk et al., 2015. van Wijk, B. C. M., Jha, A., Penny, W., and Litvak, V. (2015). Parametric estimation of cross-frequency coupling. *Journal of Neuroscience Methods*, 243:94–102.

Voytek et al., 2013. Voytek, B., D'Esposito, M., Crone, N., and Knight, R. T. (2013). A method for event-related phase/amplitude coupling. *NeuroImage*, 64(C):416–424.

Wild et al., 2010. Wild, C. J., Davis, M. H., and Johnsrude, I. S. (2010). Human auditory cortex is sensitive to the perceived clarity of speech. *NeuroImage*, 60(2):1–13.

Wild et al., 2012. Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012). Effortful Listening: The Processing of Degraded Speech Depends Critically on Attention. *Journal of Neuroscience*, 32(40):14010–14021.

7 Supporting information

S1 Fig.

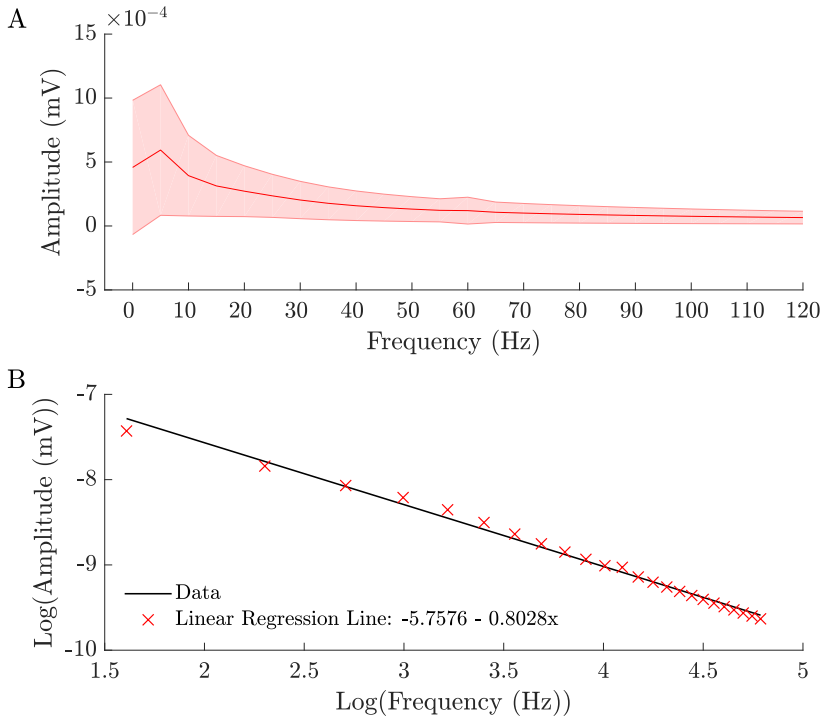


Fig S1. (A) Power vs. frequency of neural signal and (B) regression line fit. The amplitude of the neural recording signal peaks in the α (7.5-12.5 Hz) frequency range and shows the typical 1/f decay pattern thereafter. The signal depicted here is from stimulus onset (time=0) for presentation of the intact stimulus (VOC), averaged across all variables. Standard deviation across all variables is depicted by the red shaded region. The regression line fit to the neural signal confirms a 1/f decay pattern (B). The δ frequency band (0-2.5 Hz) was excluded from analyses (see Methods and Materials).

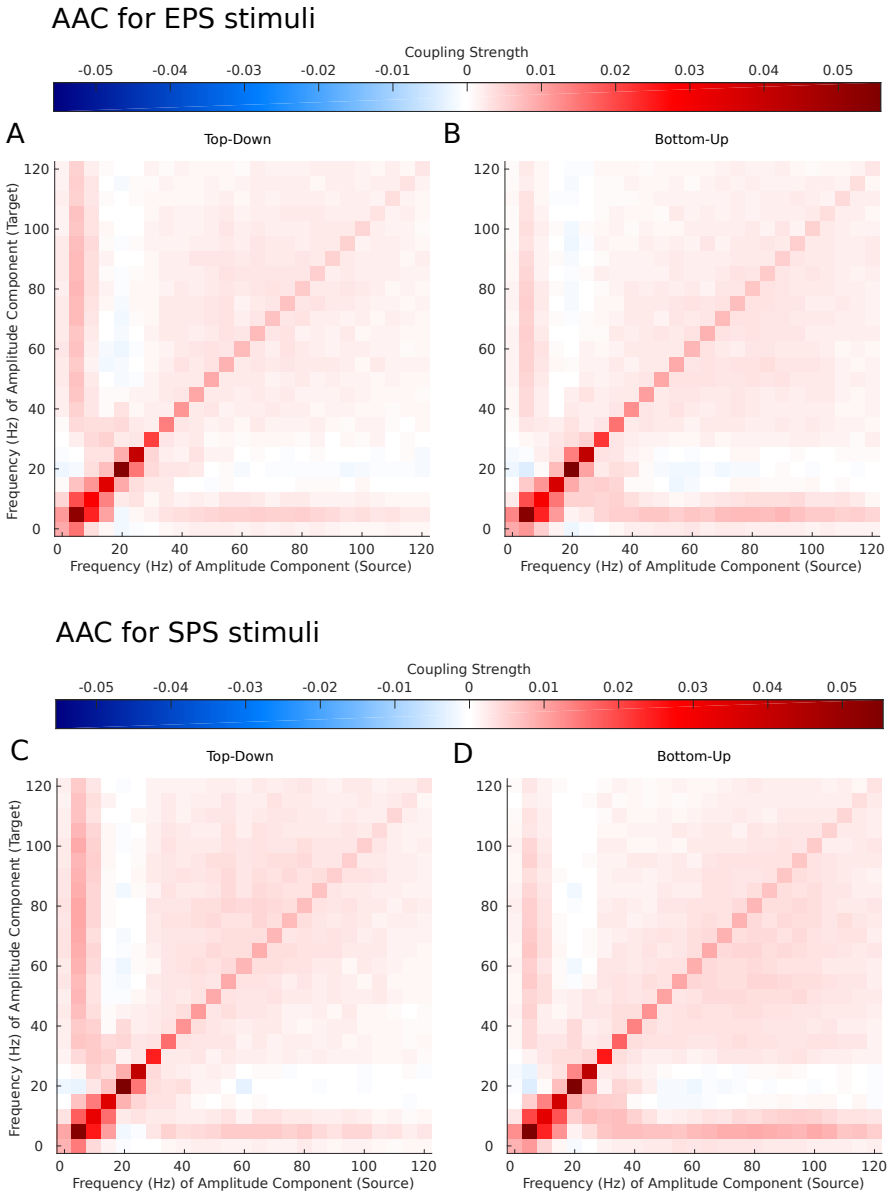


Fig S2. Top-down and bottom-up *amplitude-amplitude* coupling (AAC) strength for synthetic sounds (A-B) Top-down (A) and bottom-up (B) coupling strength in envelope-preserved sounds (EPS). Coupling strength is most pronounced in the low frequency space (<40Hz) in both directions. Depicted are canonical correlation-derived coupling coefficients, averaged across all channels, animals and cross-regional pairs.**(C-D)** Top-down (C) and bottom-up (D) coupling strength in spectrum-preserved sounds (SPS). As in EPS (A-B), coupling strength in SPS stimuli is most pronounced in the low frequency space (<40Hz). Depicted are canonical correlation-derived coupling coefficients, averaged across all channels, animals and cross-regional pairs.