

MEG Intersubject Phase Locking of Stimulus-Driven Activity during Naturalistic Speech Listening Correlates with Musical Training

Sebastian Puschmann,^{1,2,3,4} Mor Regev,¹ Sylvain Baillet,^{1,2} and Robert J. Zatorre^{1,2,5}

¹Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 2B4, Canada, ²Centre for Research on Brain, Language and Music, Montreal, Quebec H3G 2A8, Canada, ³Department of Psychology, University of Lübeck, 23562 Lübeck, Germany, ⁴Department of Psychology, Carl von Ossietzky University of Oldenburg, 26129 Oldenburg, Germany, and ⁵International Laboratory for Brain, Music and Sound Research, Montreal, Quebec H3C 3J7, Canada

Musical training is associated with increased structural and functional connectivity between auditory sensory areas and higher-order brain networks involved in speech and motor processing. Whether such changed connectivity patterns facilitate the cortical propagation of speech information in musicians remains poorly understood. We here used magnetoencephalography (MEG) source imaging and a novel seed-based intersubject phase-locking approach to investigate the effects of musical training on the interregional synchronization of stimulus-driven neural responses during listening to naturalistic continuous speech presented in silence. MEG data were obtained from 20 young human subjects (both sexes) with different degrees of musical training. Our data show robust bilateral patterns of stimulus-driven interregional phase synchronization between auditory cortex and frontotemporal brain regions previously associated with speech processing. Stimulus-driven phase locking was maximal in the delta band, but was also observed in the theta and alpha bands. The individual duration of musical training was positively associated with the magnitude of stimulus-driven alpha-band phase locking between auditory cortex and parts of the dorsal and ventral auditory processing streams. These findings provide evidence for a positive relationship between musical training and the propagation of speech-related information between auditory sensory areas and higher-order processing networks, even when speech is presented in silence. We suggest that the increased synchronization of higher-order cortical regions to auditory cortex may contribute to the previously described musician advantage in processing speech in background noise.

Key words: intersubject analysis; magnetoencephalography; musical training; phase-locking value; speech processing

Significance Statement

Musical training has been associated with widespread structural and functional brain plasticity. It has been suggested that these changes benefit the production and perception of music but can also translate to other domains of auditory processing, such as speech. We developed a new magnetoencephalography intersubject analysis approach to study the cortical synchronization of stimulus-driven neural responses during the perception of continuous natural speech and its relationship to individual musical training. Our results provide evidence that musical training is associated with higher synchronization of stimulus-driven activity between brain regions involved in early auditory sensory and higher-order processing. We suggest that the increased synchronized propagation of speech information may contribute to the previously described musician advantage in processing speech in background noise.

Received Apr. 20, 2020; revised Nov. 13, 2020; accepted Nov. 17, 2020.

Author contributions: S.P., M.R., and R.J.Z. designed research; S.P. performed research; S.P., M.R., and S.B. analyzed data; S.P. and R.J.Z. wrote the paper.

This work was supported by a research scholarship of the Deutsche Forschungsgemeinschaft (PU590/1) to S.P., a Foundation Grant from the Canadian Institutes of Health Research (Grant FDN1432179) to R.J.Z., a Discovery Grant from the Natural Science and Engineering Research Council of Canada (Grant 436355-13), the National Institutes of Health (Grant 2R01-EB-009048-05), the Healthy Brains for Healthy Lives Canada Excellence Research Fund, and a Platform Support Grant from the Brain Canada Foundation (Grant PSG15-

3755) to S.B. R.J.Z. is a fellow of the Canadian Institute for Advanced Research. We thank the members of the Research Group Auditory Cognition at the University of Lübeck for helpful comments on the data analysis and the presentation of results.

The authors declare no competing financial interests.

Correspondence should be addressed to Sebastian Puschmann at sebastian.puschmann@uol.de.

<https://doi.org/10.1523/JNEUROSCI.0932-20.2020>

Copyright © 2021 the authors

Introduction

There is increasing evidence that musical training can facilitate the perception of speech in background noise (Patel, 2014; Coffey et al., 2017a). It is assumed that this musician advantage in speech processing is, in particular, related to a superior low-level sensory processing of spoken information. In line with this view, musical training has been associated with more robust and distinct neural representations of speech in early auditory regions both in background noise and silence (Parbery-Clark et al., 2012; Bidelman et al., 2014; Coffey et al., 2017b). It is also related to enhanced processing of pitch and timing cues embedded in speech (Schön et al., 2004; Wong et al., 2007; Chobert et al., 2014; Coffey et al., 2016; Zhao and Kuhl, 2016). The potential role of changes in the higher-order processing of speech related to musical training has not been studied as much, although it has been shown that that recruitment of motor and higher-order cortical regions contributes to speech-in-noise success in musicians (Du and Zatorre, 2017).

Functional enhancements in auditory processing are accompanied by changes in cortical neuroanatomy and connectivity in musicians (Herholz and Zatorre, 2012; Moore et al., 2014; Schlaug, 2015). In the context of speech, diffusion tensor imaging data demonstrate musical training-related increases in volume and fractional anisotropy of the arcuate fasciculus (Bengtsson et al., 2005; Halwani et al., 2011; de Manzano and Ullén, 2018), which connects auditory sensory areas dorsally to the cortical speech-processing network (Friederici, 2015). Corroborating this finding, resting-state data suggest that musical training is associated with increased functional connectivity between brain regions involved in auditory sensory and speech processing (Fauvel et al., 2014; Klein et al., 2016; Palomar-Garcia et al., 2017). Whether such training-related modulations in structural and resting-state connectivity facilitate the neural spread of speech information across brain regions in musicians remains unknown.

We here investigated this question using magnetoencephalography (MEG) source imaging and an intersubject phase-locking analysis. Intersubject approaches allow the isolation of activation time series reflecting exclusively the stimulus-locked and commonly shared brain activity across a group of individuals (Hasson et al., 2004, 2010). Thus, they can reveal global patterns of stimulus- and experience-driven brain activity in naturalistic viewing or listening settings, which are also sensitive to experimental manipulations (Honey et al., 2012; Bacha-Trams et al., 2017; Regev et al., 2019) and individual differences in brain function (Naci et al., 2017). Intersubject approaches were first introduced in functional magnetic resonance imaging (fMRI), but have been extended to electrophysiological data (Dmochowski et al., 2012; Chang et al., 2015; Ki et al., 2016; Lankinen et al., 2018). Moreover, recent work demonstrates that intersubject analyses can also reveal synchronized activity between different brain areas during exposure to naturalistic stimuli, thus isolating stimulus-driven components of fMRI functional connectivity (Simony et al., 2016; Rosenthal et al., 2017; Regev et al., 2019). Here, we adapted this intersubject approach for the analysis of MEG data by computing patterns of phase synchronization across brain regions that are driven by the stimulus-locked processing of the speech input. In this analysis framework, interregional phase synchronization of MEG activity indicates that neural populations share time-locked stimulus-related information, either directly (i.e., in terms of functional connectivity) or indirectly.

We here analyzed MEG data obtained from 20 young participants with varying duration of musical training as they listened to segments of continuous natural speech. If musical training is

related to an increased neural spread of speech information, stimulus-driven phase synchronization between auditory cortex and higher-order speech-processing areas should increase as a function of training duration. Given that neural activity in auditory cortex preferentially tracks temporal fluctuations in the envelope of speech input (Ding and Simon, 2012; Kubanek et al., 2013), we expected stimulus-driven phase synchronization between auditory cortex and other brain regions to occur primarily at low frequencies, covering the delta- to alpha-band range (Ding et al., 2017; Donhauser and Baillet, 2020).

Materials and Methods

Note that all experimental data were originally collected for the study of Puschmann et al. (2019a). That study, however, did not report any of the data on speech in silence that are presented here.

Participants. Twenty right-handed volunteers (11 female; mean age: 21 ± 3 years; age range: 19–27 years) participated in the experiment. All participants were native speakers of English and had no history of neurologic, psychiatric, or hearing-related disorders. The experimental procedures were approved by the Research Ethics Board of the Montreal Neurologic Institute and Hospital, and written informed consent was obtained from all participants.

Musical training assessment. Musical training was assessed using the Montreal Music History questionnaire (Coffey et al., 2011). Data were analyzed as a function of training duration, which was defined as the number of years in which participants practiced at least three times per week for at least 1 h (Puschmann et al., 2019a). In this participant sample, the duration of musical training ranged from 0 (i.e., no musical training) to 18 years (Fig. 1C). In individuals who had received musical training, training duration was correlated with the age at training onset (age range: 5–14 years; Pearson's $r = -0.78$, $p < 0.001$) and the total number of self-reported training hours (range: 400–30,400; $r = 0.58$, $p = 0.019$). Four of 20 participants reported having received no musical training.

Experimental design. Cortical responses to continuous natural speech were recorded within the framework of a naturalistic selective listening experiment. Participants listened attentively to a continuous speech stream and detected rare target words by pressing a button. To motivate participants to listen closely to the story content instead of concentrating solely on target word detection, they were informed that the story details would have to be recalled after the experiment.

Single-speaker blocks, in which only one to-be-attended speech stream was present, alternated with selective-listening blocks, in which a second to-be-ignored speech stream was superimposed on the to-be-attended stream. Results on the selective listening task have been previously reported (Puschmann et al., 2019a); here we report exclusively on the single-speaker condition of the experiment. In total, participants were presented with 30 single-speaker blocks of 30 s duration each.

An audio recording of a detective story by Sir Arthur Conan Doyle served as natural speech input. The audio stream was cut into consecutive 30 s intervals (every other interval was used as a single-speaker block) and tagged with a sinusoidal amplitude modulation of 105 Hz modulation rate and 50% modulation depth (for an analysis of the envelope following responses; Puschmann et al., 2019a). The speech stream was presented diotically at a level of 70 dB(A). The mean word rate within stimulus blocks was 3.4 ± 0.2 Hz (range: 2.8–3.9 Hz); the mean syllable rate was 4.5 ± 0.3 Hz (range: 3.7–4.9 Hz). The mean phoneme rate was 11.2 ± 0.6 Hz (range: 9.8–12.8 Hz), as identified using WebMAUS Basic (RRID:SCR_017436; Schiel, 1999; Kislner et al., 2017).

To control for the allocation of attention toward the speech input, a different target word was displayed centrally on a presentation screen for all but the first and last 5 s of each block. For the remaining time, a fixation cross was shown to stabilize eye gaze. Target word onset lay effectively between 9.9 and 20.8 s after block onset (mean: 14.9 ± 3.5 s). In half of the blocks, the to-be-detected target word was, however, not part of the speech stream. Participants were asked to report whether the

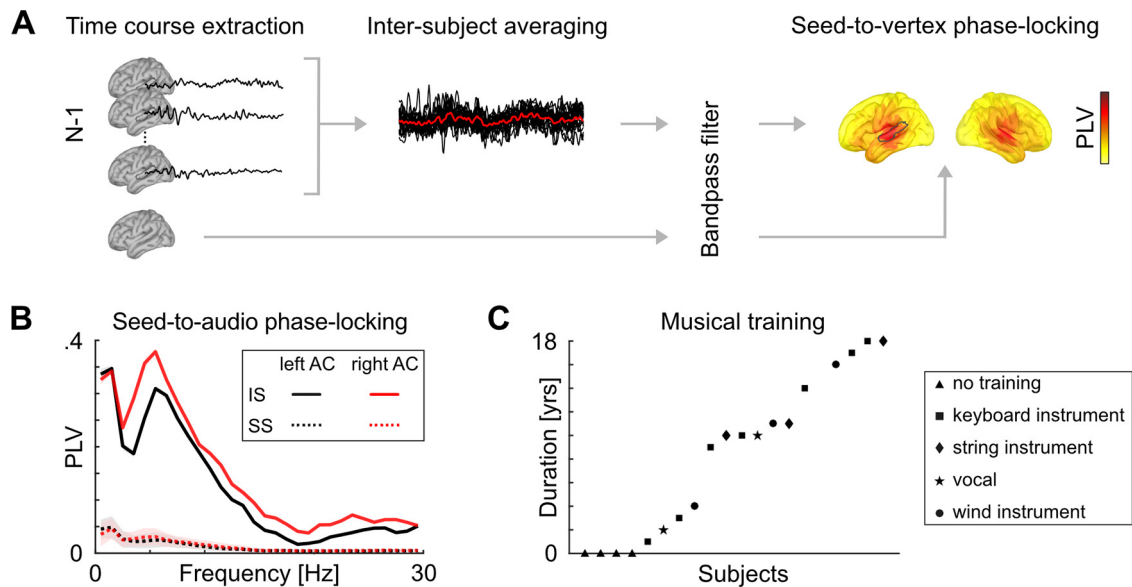


Figure 1. Intersubject phase-locking approach. **A**, Left and right auditory cortices (ACs) served as seed regions for the phase-locking analysis. For each trial, mean signal time courses were extracted from the seed region (outline shown in gray) for all but one subject. Signals were z-normalized, and averaged across subjects (black lines show individual data; the intersubject average is depicted in bold red). Phase locking was computed between the intersubject seed time course and 15,000 source time courses of the left-out subject in four frequency bands, covering the delta to beta range. For each frequency band, the resulting intersubject PLV maps were averaged across all trials and downsampled to 500 similarly sized surface parcels. This procedure was iterated for all individuals. **B**, Relative to single-subject (SS) auditory cortex time courses, intersubject (IS) averages showed increased phase locking to the envelope of the speech input, suggesting that the approach can be used to extract stimulus-driven auditory cortex activity during continuous listening. **C**, Intersubject phase locking was analyzed as a function of individual musical training. The figure depicts individual training duration and the main type of training.

target word appeared in the speech stream by pressing a button using their right hand.

Five minutes of brain activity during rest were recorded from each participant before the experiment as an empirical baseline measure of intersubject functional connectivity in the absence of auditory sensory input. During the resting-state measurement, participants were instructed to keep their eyes open and to fixate on a cross presented on the screen.

MEG data acquisition. MEG data were acquired using a 275-channel whole-head MEG system (CTF MEG). Data were recorded with a sampling rate of 2400 Hz, an antialiasing filter with a 600 Hz cutoff, and third-order spatial gradient noise cancellation. Horizontal and vertical electro-oculograms, and an electrocardiogram were acquired with bipolar montages. The head position inside the MEG sensor helmet was determined with coils fixated at the nasion and the preauricular points (fiducial points). For coregistration with anatomic MRI data, the spatial positions of the fiducial coils and of ~ 150 scalp points were obtained using a 3-D digitizer system (Isotracker, Polhemus). Participants were seated in upright position in a sound-attenuated, magnetically shielded recording room. Auditory stimulation was delivered via insert earphones (E-A-RTONE 3A, 3 M). The earphones were equipped with customized prolonged air tubes (1.5 m) to increase the distance between audio transducers and MEG sensors, thus minimizing potential stimulation artifacts in the MEG signal, which was confirmed with pilot testing using a foam head.

MEG data preprocessing. MEG data preprocessing and analysis were performed with Brainstorm (Tadel et al., 2011; RRID:SCR_001761). Line noise artifacts were removed using notch filters at the power frequency (60 Hz) and its first three harmonics. Artifacts related to eye movements and cardiac activity were pruned from the data using independent component analysis. For this procedure, a copy of the data was filtered offline between 1 and 40 Hz. A principal component analysis was performed to reduce the dimensionality of the MEG data to 40 dimensions, and 40 independent components were computed using the extended infomax algorithm implemented in Brainstorm. The demixing matrix obtained from this procedure was applied to the original unfiltered MEG dataset and independent components whose sensor topography and/or time series reflected eye blinks, lateral eye movements, or cardiac activity were

manually selected and removed. The cleaned data time series were bandpass filtered (0.1–30 Hz; zero-phase and zero-delay FIR (finite impulse response) filter, 60 dB stopband attenuation) and downsampled to a 120 Hz sampling rate for further analysis.

The task data were epoched into 30 segments of 30 s duration, corresponding to the single-speaker blocks of the experiment. Likewise, the resting-state data were cut into 10 consecutive 30 s intervals.

MEG data were analyzed with cortical source modeling. The participants' individual structural T1-weighted MRI images were automatically segmented and labeled using Freesurfer (RRID:SCR_001847; Dale et al., 1999; Fischl et al., 1999, 2004), and coregistered to the MEG data in Brainstorm using the digitized head points. An OPENMEEG boundary element method head model and a minimum-norm source model with $\sim 15,000$ surface vertices and depth weighting (order: 0.5; maximal amount: 10) was computed for each participant. Dipole orientation was constrained to be normal to the cortical surface. Noise covariance matrices for the source reconstruction process were estimated from 2 min empty-room recordings obtained for each participant.

Intersubject PLV computation. The cortical spread of stimulus-driven neural activity during listening to continuous speech was quantified in terms of phase synchronization of source-localized MEG activity to the intersubject time courses of auditory cortex activity. The phase-locking value (PLV), which represents the absolute value of the mean phase difference between two time series expressed as a complex unit-length vector, served as a metric of phase synchronization in each listening block of the experiment (Lachaux et al., 1999; Mormann et al., 2000).

Figure 1A depicts the pipeline of the intersubject PLV analysis. For each participant and each trial, mean source time courses were extracted from left and right auditory cortex seed regions. These seed regions corresponded to previously identified parts of the temporal lobe that showed robust tracking of the speech envelope (Puschmann et al., 2019a). They encompassed Heschl's gyrus, the posterior planum polare, planum temporale, and adjacent parts of the superior temporal gyrus. The extracted seed time courses were z-normalized to account for individual differences in overall signal amplitude and averaged across all but one participant iteratively. Phase locking was computed between this intersubject seed time course and all 15,000 source time courses of the left-out participant, resulting in a whole-brain intersubject PLV map.

Note that the intersubject averaging of the seed time courses and the leave-one-out cross-validation approach ensured that phase locking was driven only by the commonly shared and time-locked processing of the speech stimulus in auditory cortex and not by individual intrinsic oscillatory activity. Complex PLV estimates were subsequently averaged across all 30 trials, PLV magnitude was extracted, and maps were projected to MNI stereotaxic space and downsampled to a random atlas parcellation that consisted of 500 similarly sized surface parcels. Downsampling aimed to reduce the number of statistical comparisons while retaining adequate spatial resolution. The random atlas parcellation was generated using the surface clustering tool (homogeneous parcellation: random) provided by Brainstorm. In detail, left and right hemisphere surfaces were generated by merging respective areas of the Desikan-Killiany atlas template. Each surface was then parcellated in a stepwise manner into 10 and 25 homogeneous surface parcels.

Previous work suggests that neural activity in auditory cortex tracks slow temporal modulations of speech, such as the speech envelope (e.g., Kubanek et al., 2013). Slow temporal modulations in speech occur in the delta to alpha range and are maximal in the theta band (Ding et al., 2017). We therefore assumed that intersubject synchronization of auditory cortex activity and intersubject phase locking predominantly occurs in this frequency range. Previous MEG studies further suggest frequency-specific interactions between auditory cortex and higher-order brain regions during speech processing (Park et al., 2015; Keitel et al., 2017, 2018), potentially reflecting the processing of phonemic and prosodic features. We therefore further subdivided the delta to alpha range into three frequency bands for the whole-brain PLV analysis, covering the delta (1–4 Hz), theta (4–8 Hz), and alpha (8–12 Hz) bands. Although we did not expect robust intersubject synchronization and phase locking at higher frequencies, we also included the beta band (15–30 Hz) as a control. Frequency bands were extracted automatically by Brainstorm during PLV computation using an FFT-based FIR filter with no phase shift (Brainstorm function: `bst_bandpass_fft`).

To explore the spectral profile of the observed phase-locking pattern in more detail, we computed phase locking within 30 narrow frequency bands that uniformly covered the frequency range from 0.1 to 30 Hz, and for different surface parcels localized to auditory cortex and temporal lobe, inferior and middle frontal gyrus, the premotor area, and inferior parietal lobe, all of which showed robust stimulus-driven phase synchronization to auditory cortex activity in the delta- to alpha-band range. As all tested surface parcels showed a similar spectral PLV profile, we exemplarily depict data of only four regions in the article: left and right auditory cortex, left inferior frontal lobe, and right precentral gyrus.

Previous fMRI studies confirmed that no meaningful intersubject alignment of signal time courses and no intersubject functional correlation between brain regions could be observed in the absence of a driving external sensory stimulation (Simony et al., 2016). Hence, we here computed intersubject PLV obtained from the 5 min resting-state recording as an empirical baseline for the analysis of stimulus-driven functional connectivity. The analysis pipeline for the resting state was identical to the task data.

To control whether intersubject averaging of source time courses, as expected, strengthens the temporal relationship between ongoing auditory cortex activity and the actual sound input, we computed phase locking between the envelope of the audio input and the 20 averaged auditory cortex time courses obtained using the leave-one-out approach described above (Fig. 1B). The resulting phase-locking pattern was compared with PLVs obtained for the audio envelope and the 20 individual (i.e., single-subject) auditory cortex time courses. The speech envelope was determined as described by Puschmann et al. (2019b). Audio signals were *z*-normalized and bandpass filtered into 128 logarithmically spaced frequency bands between 100 and 6500 Hz with a gammatone filter bank (Hohmann, 2002; Hercke and Hohmann, 2004). A Hilbert transformation was used to compute the signal envelope within each of the 128 frequency bands. The envelope was then obtained by averaging the absolute Hilbert values across all bands. PLV was computed for each trial and 30 frequency bands, uniformly covering the frequency range from 0.1 to 30 Hz (for filter details, see above). The obtained PLV profiles were subsequently averaged across the 30 trials.

A potential issue arising from MEG source imaging is that cross talk, which is a spatial leakage phenomenon between source activity reconstructed at proximate brain locations, may confound data interpretation. Although the present intersubject phase-locking approach should be relatively insensitive to such adverse effects when mapping individual spontaneous activity that is not phase locked to the speech stimulus, the issue could manifest when detecting auditory cortex activity evoked by the same acoustic stimulus across individuals, potentially contributing to the observed phase-locking patterns in nearby areas. We therefore conducted simulations using synthesized data to investigate the potential confounding influence of source signal leakage from auditory cortex activity.

This analysis was performed using the Brainstorm data simulation toolkit. We used the same individual head models, and sensor and environmental noise statistics as the original data to synthesize MEG sensor time series generated from bilateral auditory cortex activity. Sensor noise levels were computed from the individual noise covariance matrices, which were based on empty-room MEG recordings. The source locations were the brain regions used as seeds for the intersubject analysis. We then mapped the resulting simulated sensor data onto individual brains to assess the extent of source signal leakage.

We used the envelope of the speech signal (for details on envelope extraction, see above) plus random noise as the driving source time series in the seed auditory cortex regions. The mean resulting signal-to-noise ratio (SNR) of this ground truth signal was -3.4 dB [defined as $\text{rms}_{\text{Signal}}/(\text{rms}_{\text{Signal}} + \text{rms}_{\text{Noise}})$]. Note that the SNR and therefore the to-be-expected signal leakage of the simulated data were higher than for the actual MEG task data (mean SNR = -9.3 dB, based on the amplitude of the intersubject average relative to the amplitude of the total signal obtained from the auditory cortex seed). We ran signal leakage simulations for each subject and all 30 trials of the experiment.

We derived the same phase-locking metric (i.e., PLV) between the resulting source time series across the entire cortex and the speech signal envelope within four different frequency bands (1–4, 4–8, 8–12, and 15–30 Hz). Complex PLV estimates were subsequently averaged across the 30 trials, PLV magnitude was extracted, and individual maps were projected to the MNI stereotaxic space and downsampled to 500 similarly sized surface parcels for statistical analyses.

Statistical analysis. The statistical analysis of MEG and behavioral data were based on all 20 datasets. All statistical testing was performed using MATLAB (RRID:SCR_001622). Effects were deemed statistically significant at $p < 0.05$, corrected for multiple comparisons. For parcel-level MEG data analyses, *p* values were adjusted for false discovery rate (FDR) over 500 comparisons (i.e., the number of brain parcels) following Benjamini and Hochberg (1995). For all other analyses, Bonferroni correction was applied.

Spatial patterns of brain regions showing a robust phase synchronization to auditory cortex seed time series were assessed by computing paired *t* tests between the respective intersubject PLV maps obtained for listening and resting state. As a reduction of stimulus-driven phase synchronization during listening relative to the resting state baseline cannot be interpreted meaningfully, one-tailed testing was performed.

Intersubject PLV patterns and simulated PLV maps generated for a bilateral auditory cortex source were compared using two-tailed paired *t* tests to control for effects of signal leakage on intersubject phase locking. Individual PLV maps were *z*-transformed before *t* tests were performed to account for overall differences in PLV magnitude between simulation and real task data.

Spearman correlation was used to investigate the relationship between the duration of musical training and intersubject PLV maps obtained from the listening blocks. Based on our hypothesis that musical training strengthens the stimulus-driven phase synchronization between auditory cortex and higher-order brain regions involved in speech processing, we tested only for positive correlations. Statistical testing was performed on the parcel level, and separately for both seeds (i.e., left and right auditory cortex) and the four frequency bands of interest.

Qualitatively, the effects of musical training on intersubject PLV from left and right seed regions were not symmetric. As a *post hoc* test of lateralization, we determined the laterality index (LI) based on the

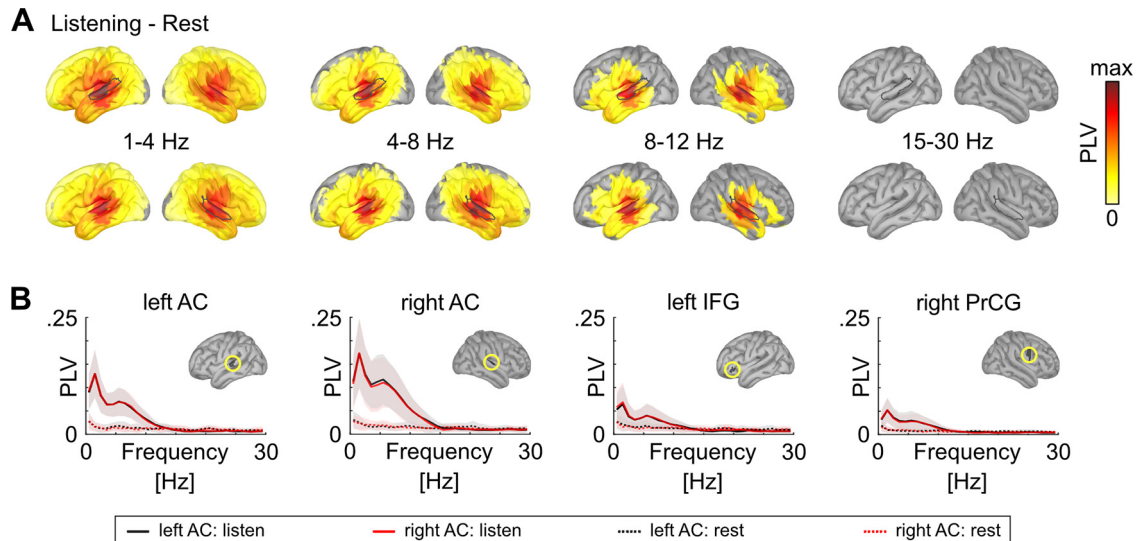


Figure 2. Phase synchronization to stimulus-driven auditory cortex (AC) activity during continuous naturalistic listening. **A**, Relative to resting state, we observed enhanced stimulus-driven phase synchronization in the delta (1–4 Hz), theta (4–8 Hz), and alpha (8–12 Hz) range while listening to the speech input (at $p < 0.05$, FDR corrected for multiple comparisons). For all of these frequency bands and both seed regions (top row: left auditory cortex; bottom row: right auditory cortex; seed outline shown in gray), phase locking was maximal in ipsilateral and contralateral auditory sensory areas. Phase-locking magnitude decreased with increasing frequency; the color bar is adjusted to the maximum (max) phase-locking value in each frequency band. **B**, To investigate the spectral profile of the observed intersubject phase locking in more detail, PLV was computed for 30 consecutive frequency bands (0.1–30 Hz) between the auditory cortex seed time courses and single surface parcels localized to temporal, frontal, and inferior parietal cortices. For all tested brain regions, a consistent PLV profile was observed, which suggests that intersubject phase locking effectively shows two distinct peaks in the delta- and theta-band range. Exemplarily, the figure depicts the PLV pattern during listening (solid lines) and resting state (dashed lines) for left AC, right AC, left inferior frontal gyrus (IFG), and right precentral gyrus (PrCG).

number of surface parcels showing significant positive correlations within each cortex hemisphere. LI was computed as $(N_{\text{Left}} - N_{\text{Right}}) / (N_{\text{Left}} + N_{\text{Right}})$, with N_{Left} and N_{Right} corresponding to the number of parcels in the left or right hemisphere in which a significant relationship between PLV and the duration of musical training was observed (Seghier, 2008). Lateralization was determined based on the mean LI obtained at four different statistical thresholds ($p = 0.05, 0.01, 0.005, 0.001$, uncorrected) to minimize effects of the used significance criterion. The p values corresponding to a given LI were estimated using a resampling approach (10,000 iterations) in which hemisphere labels (left, right) were randomly permuted across brain regions to achieve a sampling distribution of the test statistic under the null hypothesis.

To control for effects of musical training on behavioral performance, potentially reflecting differences in task engagement or attention allocation, two-tailed Spearman correlations between the duration of musical training and individual hit rates, false alarm rates, as well as response times in the target word detection task were computed. Button presses that were registered between 0 and 3000 ms after the onset of the target word were regarded as hits; button presses occurring outside of these intervals were treated as false alarms.

Results

Behavioral data

To control for attention to the continuous speech input, participants were asked to respond to rare target words. Overall, $93.0 \pm 7.6\%$ of the target words (mean \pm SD; range: 73.3–100%) were detected; $6.3 \pm 9.2\%$ (range: 0–25%) of button presses were false alarms. The mean response time to target words was 730 ± 134 ms (range: 555–1035 ms). None of these behavioral measures was correlated with the duration of musical training (hit rate: $\rho = 0.04$, $p = 0.883$; false alarm rate: $\rho = 0.18$, $p = 0.458$; response times: $\rho = -0.01$, $p = 0.957$), indicating that the effects of musical training on neural processing are unlikely to reflect increased task engagement in this simple listening situation in which performance is near ceiling.

Spatial pattern of stimulus-driven phase synchronization

To explore the extent of neural synchronization along the auditory-processing hierarchy during speech listening, we computed intersubject phase locking between the auditory cortices and the rest of the brain. Figure 2A depicts brain areas in which phase locking was significantly increased during listening compared with resting state (at $p < 0.05$, FDR corrected), for left (Fig. 2A, top) and right (Fig. 2A, bottom) auditory cortex seeds and all tested frequency bands. Stimulus-driven intersubject phase locking was observed in the delta (i.e., 1–4 Hz), theta (4–8 Hz), and alpha (8–12 Hz) ranges, but not for the beta range. PLV patterns were largely symmetrical, with global maxima localized over ipsilateral and contralateral auditory sensory regions. Phase-locking magnitude generally decreased with increasing band frequency; the overall PLV pattern, however, remained largely unchanged. As expected without a common driving sensory input, no intersubject phase synchronization was observed during resting state, which was used as an empirical baseline condition for this comparison (Fig. 2B, dotted lines; whole-brain maps not shown).

To study in more detail how phase-locking patterns evolve with frequency, we computed intersubject phase locking for different auditory and frontal brain regions over 30 narrow frequency bands across the [0.1, 30] Hz range. The resulting spectral intersubject PLV profile of all tested brain regions effectively showed two distinct peaks in the delta range (in the 1–2 Hz band) and the theta range (5–6 Hz band), but no peak in the alpha band. Figure 2B depicts the spectral intersubject PLV profile for left and right auditory cortex, left inferior frontal gyrus, and right precentral gyrus. Of note, a similar PLV profile was also found when computing phase locking between the intersubject seed time courses and the envelope of the speech stimulus (Fig. 1B). This suggests that the observed phase-locking profile reflects the temporal structure of the auditory input itself.

We ran simulations to investigate the potential confounding influence of source signal leakage from auditory cortex activity

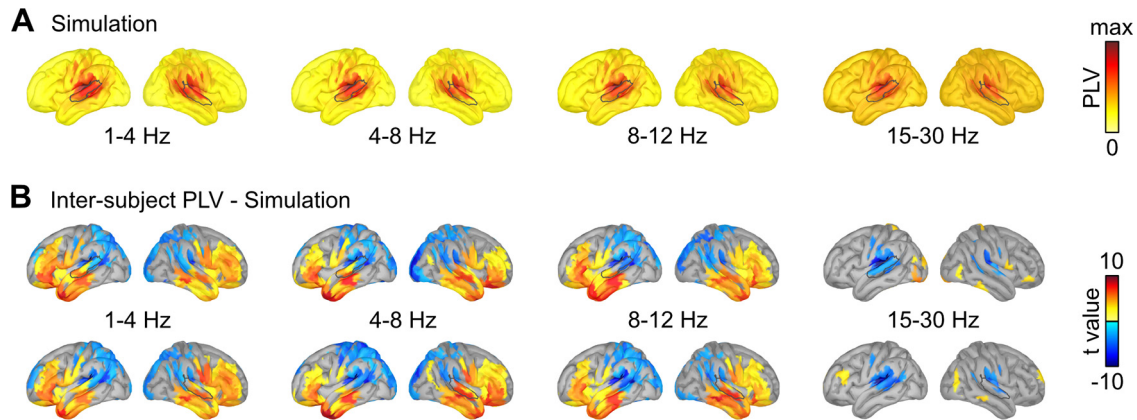


Figure 3. Simulation of phase locking expected as a result of signal leakage from a bilateral auditory source (outline shown in black). **A**, Simulated stimulus-driven phase locking was maximal over auditory cortices, spreading toward adjacent brain regions. Phase-locking magnitude decreased from the delta to the beta band; the color bar is adjusted to the maximum (max) phase-locking value in each frequency band. **B**, Compared with the original data, simulated phase locking was significantly reduced over frontal and anterior temporal regions, whereas phase locking to parietal cortex was relatively enhanced (two-tailed paired t test at $p < 0.05$, FDR corrected).

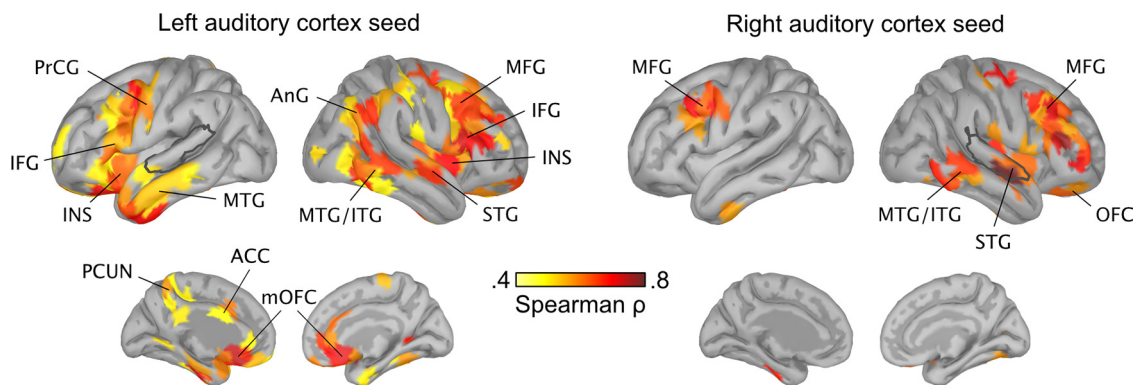


Figure 4. Relationship between stimulus-driven phase synchronization during listening to continuous speech and musical training: Spearman correlation (one-tailed) showed a positive association between the duration of individual musical training and the magnitude of alpha-band phase locking between auditory cortex seed time courses and various brain regions in frontal and temporal cortices, and along the midline (at $p < 0.05$, FDR correction). Brain regions for which stimulus-driven interregional phase synchronization was positively associated with musical training included the superior (STG), middle (MTG) and inferior temporal gyri (ITG); the insula (INS); the orbitofrontal (OFC) and medial orbitofrontal (mOFC) cortices; the inferior (IFG) and middle frontal (MFG) gyri; the precentral (PrCG) gyrus; the angular gyrus (AnG); the precuneus (PCUN); and the anterior cingulate cortex (ACC).

on the observed intersubject phase-locking patterns. **Figure 3A** depicts phase locking between the envelope of the speech input and synthesized source time courses of a bilateral auditory cortex source that tracks fluctuations in the sound envelope. The phase-locking map from the synthesized data peaked in the ground truth auditory cortex regions, spreading to adjacent brain regions. The phase-locking magnitude decreased from delta to beta band, which to some extent resembled the PLV patterns observed in our original data. However, synthesized phase locking was significantly reduced over frontal and anterior temporal regions, compared with the original data (**Fig. 3B**; paired t test, $p < 0.05$, FDR corrected; note that PLV was higher over parietal regions in the synthesized data). We conclude from this simulation that the patterns of phase synchronization between auditory cortex and frontal and temporal regions, which were salient in the original data, are unlikely to be caused by stimulus-driven signal leakage from auditory cortex. The data instead provide evidence for active frontal and anterior temporal processes resulting in brain activity synchronized in phase with auditory cortex activations. The reduced intersubject phase locking in parietal regions compared with synthesized patterns may, in turn, point to the presence of parietal activations in the original data that are not phase locked to the sound envelope reference.

However, note that the SNR of the simulated data were higher than for the original data, potentially resulting in an increased signal spread toward the parietal lobe.

Musical training and stimulus-driven phase synchronization

Based on previous data on structural and resting-state connectivity, we hypothesized that musical training should be associated with an increased neural spread of speech information across the auditory processing hierarchy. To investigate this question, the magnitude of stimulus-driven phase locking was analyzed as a function of musical training duration. Our data show positive relationships between the individual duration of musical training and the magnitude of intersubject phase locking between the auditory cortex seeds and widespread frontotemporal as well as parietal brain regions, including the middle temporal gyrus, the temporal pole region, the inferior and middle frontal gyri, the precentral and postcentral gyri, the angular gyrus, the anterior and posterior cingulate cortices, the medial prefrontal cortex, the orbitofrontal cortex, and the precuneus (**Fig. 4**; one-tailed Spearman correlation, $p < 0.05$, FDR corrected).

Effects of musical training on cortical phase synchronization of neural activity could be observed for left and right auditory cortex seeds. Qualitatively, musical training was more associated

with the spread of speech information from the left auditory cortex. As predicted, the observed anatomic pattern was qualitatively similar to the structure of the dual-stream model of speech processing, including both ventral and dorsal cortical regions (Hickok and Poeppel, 2007). Effects of musical training, however, extended beyond this speech-processing network and were not lateralized toward the left hemisphere. Instead, musical training was also related to increased phase locking to homologous speech-processing areas in the right hemisphere. Overall, the data even suggest some degree of rightward lateralization, both for the left ($LI = -0.22$; $p = 0.014$, Bonferroni corrected) and the right auditory cortex seed ($LI = -0.48$; $p < 0.001$, Bonferroni corrected).

The reported effects of musical training were exclusive to the alpha band (i.e., 8–12 Hz), which represents the upper boundary of frequencies in which stimulus-driven phase locking was observed (Fig. 2, compare *A*, *B*). Except for a single surface parcel in the right frontal pole showing a positive association between delta-band phase locking and musical training duration, no effects of musical training were found for any of the other frequency bands (at $p < 0.05$, FDR corrected). As an exploratory analysis, we also checked for potential negative correlations between intersubject PLV and the duration of musical training. This analysis did not reveal statistically significant effects in any of the tested frequency bands (one-tailed Spearman correlation, all $p > 0.4$, FDR corrected).

Discussion

We here investigated the effect of musical training on the neural synchronization between auditory cortices and high-order brain regions during continuous listening to naturalistic speech. For this purpose, we implemented a novel intersubject phase-locking method that allowed isolating patterns of interregional phase synchronization driven by the commonly shared processing of speech signals in auditory cortex (Fig. 1). Our results show robust cortical patterns of phase synchronization in the delta-to-alpha-band range, in line with the modulation spectrum of continuous speech (Ding et al., 2017). Stimulus-driven phase locking was maximal in a bilateral frontoparietal brain network (Fig. 2). Across participants, stimulus-driven phase synchronization between auditory cortex and multiple frontoparietal and temporal brain structures was positively associated with the duration of musical training (Fig. 4). This finding provides evidence that musical training can facilitate the cortical spread of stimulus-driven auditory cortex activity, in line with previous anatomic and resting-state data, suggesting an increased connectivity between auditory sensory and higher-order regions involved in speech and motor processing.

Stimulus-driven phase synchronization during listening to natural speech

There exist two main cortical speech-processing streams, originating from auditory cortex: a ventral stream that connects auditory cortex via the superior temporal gyrus to ventrolateral frontal cortex; and a dorsal stream that connects posterior portions of auditory cortex to premotor cortex via the inferior parietal lobe (Hickok and Poeppel, 2007; Petrides and Pandya, 2009; Rauschecker and Scott, 2009; Friederici, 2011). Our data show widespread phase synchronization between stimulus-driven activity in auditory cortex and neural response time courses in several brain regions that are part of both streams, including

superior and middle temporal gyrus, inferior parietal lobe, premotor cortex, and inferior frontal regions. While previous work provided compelling evidence that speech-processing streams are lateralized toward the left hemisphere (Vigneau et al., 2011; Hickok, 2012), our current data, however, show largely symmetrical bilateral patterns of stimulus-driven phase synchronization. We assume that this pattern predominantly reflects the neural spread of acoustic information from early sensory brain regions, as ongoing activity in auditory cortex is tightly linked to the acoustic properties of the speech signal (Kubaneck et al., 2013; Brodbeck et al., 2018b; but see Brodbeck et al., 2018a; Donhauser and Baillet, 2020).

Increased interregional phase synchronization in individuals with musical training

Previous studies provided evidence for a more developed macrostructure and microstructure of the fiber bundles underlying the higher-order auditory processing pathways in professional musicians compared with nonmusicians (Bengtsson et al., 2005; Halwani et al., 2011; de Manzano and Ullén, 2018; Oechslin et al., 2018). These observations are presumably related to the important sensory–motor integration demands of musical performance (Zatorre et al., 2007). Complementing these anatomic data, resting-state data suggest an increased functional connectivity between auditory cortices and the cortical speech-processing network in musicians (Fauvel et al., 2014; Klein et al., 2016; Palomar-García et al., 2017). These functional and anatomic features may potentially lead to a more efficient flow of spoken information in musicians from auditory sensory to higher-order brain regions and, conversely, may also provide an enhanced pathway for top-down effects. In line with this notion, our current data show a positive association between the duration of individual musical training and the stimulus-driven information shared between auditory cortex and higher-order brain regions while listening to continuous speech. Thus, we demonstrate that the previously reported changes in neural wiring infrastructure may indeed support an increased spread of task-relevant information. The observed pattern of brain regions in which phase locking increased as a function of musical training duration were largely in line with the anatomic structure of the dual-stream model of higher-order auditory processing. Effects of musical training, however, exceeded the traditional language network and encompassed also other higher-order cognitive regions such as the precuneus, orbitofrontal areas, and the medial prefrontal cortex, which have been shown to be involved in processing of spoken stories (Xu et al., 2005; Lerner et al., 2011; Ben-Yakov et al., 2012; AbdulSabur et al., 2014).

While resembling the overall pattern of speech processing, effects of musical training were—again—not lateralized to the left hemisphere in our dataset (although the effects appeared to be stronger for the left auditory cortex seed). Instead, a longer duration of musical training was also associated with increased stimulus-driven phase synchronization between auditory cortices and brain regions in the right ventral and dorsal processing streams, effectively even leading to some degree of rightward lateralization. Neural activity along these auditory processing streams of the right hemisphere has been previously associated with the neural processing of music (Koelsch et al., 2005; Hyde et al., 2006; Klein and Zatorre, 2011). It seems likely that this rightward lateralization reflects the proposed right-hemispheric bias for fine-grained spectral processing that underlies the perception of music (Zatorre et al., 2002; Poeppel, 2003; Albouy et al., 2020). Our data suggest that musicians may show a stronger recruitment of the

dorsal and ventral streams during speech processing both in the left hemisphere, supporting more canonical linguistic analysis, as well as in the right hemisphere, which may reflect a more extensive processing of spectral cues, including fundamental frequency variations, harmonic information in vowels, vocal timbre, and other paralinguistic elements, all of which could also be relevant to segregate speech sounds from background noise (Sammler et al., 2015; Kreitewolf et al., 2018).

Effects of musical training on stimulus-driven phase synchronization were restricted to the alpha band, despite intersubject phase locking being more pronounced in the slower delta and theta ranges (Fig. 2A). In principle, the alpha range corresponds to the phoneme rate of natural speech (Keitel et al., 2018; mean phoneme rate in this study: 11.2 ± 0.6 Hz). Therefore, increased alpha-band phase locking could potentially reflect a superior phoneme-level processing (Di Liberto et al., 2015; Daube et al., 2019) in musically trained individuals. In line with this notion, there exist data showing better behavioral phoneme classification in musicians (Varnet et al., 2015; but see Swaminathan and Schellenberg, 2017). However, the spectral phase-locking profile (Fig. 2B) showed no peak in the alpha range that would be indicative of specific phoneme-level processing. Therefore, the exact functional processes underlying the observed increase of interregional neural synchronization in the alpha band, which represents the upper frequency range of slow temporal modulations in speech (Ding et al., 2017), remain unclear at present and need further analysis. Given that no increased phase synchronization related to musical training was observed in auditory cortex itself (i.e., the seed region of our analysis), it seems rather unlikely that the effect can be simply explained by a more accurate auditory cortex tracking of alpha-band modulations in speech (Fig. 4).

Musical training and speech-in-noise perception

Previous research suggests that musical training can benefit speech perception in adverse listening conditions (Coffey et al., 2017a; Puschmann et al., 2019a), and may to some extent mitigate the age-related decline in segregating target sounds from background noise (Parbery-Clark et al., 2011; Alain et al., 2014). Beneficial effects of musical training on speech processing have been associated with a superior low-level processing of auditory information (Strait et al., 2014), even when measured in quiet (Coffey et al., 2016) like in the present study, as well as with enhancements in top-down processing (Strait and Kraus, 2011; Kraus et al., 2012; Du and Zatorre, 2017). Based on our current findings, we speculate that increased synchronization of stimulus-driven neural activity between auditory cortex and linguistic or other higher-cognitive cortical areas may also contribute to this musician advantage, because it reflects higher rates of information sharing between sensory processing areas and higher-order mechanisms involved in top-down processes. However, additional work is necessary to test this hypothesis in more detail and to understand the potential role of patterns of information propagation along the processing hierarchy.

Conclusions

Intersubject comparisons have been shown to help remove intrinsic neural fluctuations and non-neural artifacts (e.g., physiological noise), and, therefore, to improve sensitivity to stimulus-locked information (Simony et al., 2016). We here demonstrated that this approach can be adapted for MEG to reveal interregional patterns of synchronized stimulus-driven brain activity. By measuring responses to complex time-varying stimuli, intersubject approaches provide a complementary method to resting-

state connectivity (Liu et al., 2010; Brookes et al., 2011; Hipp et al., 2012). It can serve as a tool to track individual differences in information propagation, for example, as a result of training or neurologic disorders, and the modulation of functional networks by cognitive processing.

Our data provide new evidence for a positive relationship between musical training and the propagation of stimulus-related information along the auditory processing hierarchy. This finding complements previous reports on strengthened anatomical and resting-state connections of auditory cortex in musicians and demonstrates that enhanced connectivity patterns may indeed facilitate the cortical processing of spoken information. Our results add to the view that musical training can benefit the processing of speech information and suggests that the previously reported musician benefit in speech-in-noise perception may not only be related to a superior low-level processing of speech information within the auditory system, but also to a stronger neural synchronization between brain regions involved in early sensory and higher-order auditory processing.

References

- AbdulSabur NY, Xu Y, Liu S, Chow HM, Baxter M, Carson J, Braun AR (2014) Neural correlates and network connectivity underlying narrative production and comprehension: a combined fMRI and PET study. *Cortex* 57:107–127.
- Alain C, Zedel BR, Hutka S, Bidelman GM (2014) Turning down the noise: the benefit of musical training on the aging auditory brain. *Hear Res* 308:162–173.
- Albouy P, Benjamin L, Morillon B, Zatorre RJ (2020) Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science* 367:1043–1047.
- Bacha-Trams M, Glerean E, Dunbar R, Lahnakoski JM, Ryyppö E, Sams M, Jääskeläinen IP (2017) Differential inter-subject correlation of brain activity when kinship is a variable in moral dilemma. *Sci Rep* 7:14244.
- Ben-Yakov A, Honey CJ, Lerner Y, Hasson U (2012) Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *Neuroimage* 63:501–506.
- Bengtsson SL, Nagy Z, Skare S, Forsman L, Forssberg H, Ullén F (2005) Extensive piano practicing has regionally specific effects on white matter development. *Nat Neurosci* 8:1148–1150.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57:289–300.
- Bidelman GM, Weiss MW, Moreno S, Alain C (2014) Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur J Neurosci* 40:2662–2673.
- Brodbeck C, Hong LE, Simon JZ (2018a) Rapid transformation from auditory to linguistic representations of continuous speech. *Curr Biol* 28:3976–3983.e5.
- Brodbeck C, Presacco A, Simon JZ (2018b) Neural source dynamics of brain responses to continuous stimuli: speech processing from acoustics to comprehension. *Neuroimage* 172:162–174.
- Brookes MJ, Woolrich M, Luckhoo H, Price D, Hale JR, Stephenson MC, Barnes GR, Smith SM, Morris PG (2011) Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proc Natl Acad Sci U S A* 108:16783–16788.
- Chang WT, Jääskeläinen IP, Belliveau JW, Huang S, Hung AY, Rossi S, Ahveninen J (2015) Combined MEG and EEG show reliable patterns of electromagnetic brain activity during natural viewing. *Neuroimage* 114:49–56.
- Chobert J, François C, Velay JL, Besson M (2014) Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cereb Cortex* 24:956–967.
- Coffey EBJ, Herholz SC, Scala S, Zatorre RJ (2011) Montreal Music History Questionnaire: a tool for the assessment of music-related experience in music cognition research. Poster presented at The Neurosciences and Music IV: Learning and Memory Conference, Edinburgh, UK.

- Coffey EB, Colagrosso EM, Lehmann A, Schönwiesner M, Zatorre RJ (2016) Individual differences in the frequency-following response: relation to pitch perception. *PLoS One* 11:e0152374.
- Coffey EBJ, Mogilever NB, Zatorre RJ (2017a) Speech-in-noise perception in musicians: a review. *Hear Res* 352:49–69.
- Coffey EBJ, Chepesiuk AMP, Herholz SC, Baillet S, Zatorre RJ (2017b) Neural correlates of early sound encoding and their relationship to speech-in-noise perception. *Front Neurosci* 11:479.
- Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* 9:179–194.
- Daube C, Ince RAA, Gross J (2019) Simple acoustic features can explain phoneme-based predictions of cortical responses to speech. *Curr Biol* 29:1924–1937.e9.
- de Manzano Ö, Ullén F (2018) Same genes, different brains: neuroanatomical differences between monozygotic twins discordant for musical training. *Cereb Cortex* 28:387–394.
- Di Liberto GM, O’Sullivan JA, Lalor EC (2015) Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Curr Biol* 25:2457–2465.
- Ding N, Simon JZ (2012) Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol* 107:78–89.
- Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppel D (2017) Temporal modulations in speech and music. *Neurosci Biobehav Rev* 81:181–187.
- Dmochowski JP, Sajda P, Dias J, Parra LC (2012) Correlated components of ongoing EEG point to emotionally laden attention - a possible marker of engagement? *Front Hum Neurosci* 6:112.
- Donhauser PW, Baillet S (2020) Two distinct neural timescales for predictive speech processing. *Neuron* 105:385–393.e9.
- Du Y, Zatorre RJ (2017) Musical training sharpens and bonds ears and tongue to hear speech better. *Proc Natl Acad Sci U S A* 114:13579–13584.
- Fauvel B, Groussard M, Chételat G, Fouquet M, Landeau B, Eustache F, Desgranges B, Platel H (2014) Morphological brain plasticity induced by musical expertise is accompanied by modulation of functional connectivity at rest. *Neuroimage* 90:179–188.
- Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9:195–207.
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically parcellating the human cerebral cortex. *Cereb Cortex* 14:11–22.
- Friederici AD (2011) The brain basis of language processing: from structure to function. *Physiol Rev* 91:1357–1392.
- Friederici AD (2015) White matter pathways for speech and language processing. In: *Handbook of clinical neurology: the human auditory system* (Ceslia GG, Hickok G, eds), pp 177–186. Amsterdam: Elsevier.
- Halwani GF, Loui P, Rüber T, Schlaug G (2011) Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Front Psychol* 2:156.
- Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. *Science* 303:1634–1640.
- Hasson U, Malach R, Heeger DJ (2010) Reliability of cortical activity during natural stimulation. *Trends Cogn Sci* 14:40–48.
- Herholz SC, Zatorre RJ (2012) Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76:486–502.
- Herzke T, Hohmann V (2004) Improved numerical methods for gammatone filterbank analysis and synthesis. *Acta Acust United Acust* 93:498–500.
- Hickok G (2012) The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. *J Commun Disord* 45:393–402.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
- Hipp JF, Hawellek DJ, Corbetta M, Siegel M, Engel AK (2012) Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat Neurosci* 15:884–890.
- Hohmann V (2002) Frequency analysis and synthesis using a Gammatone filterbank. *Acta Acust United Acust* 88:433–442.
- Honey CJ, Thompson CR, Lerner Y, Hasson U (2012) Not lost in translation: neural responses shared across languages. *J Neurosci* 32:15277–15283.
- Hyde KL, Zatorre RJ, Griffiths TD, Lerch JP, Peretz I (2006) Morphometry of the amusic brain: a two-site study. *Brain* 129:2562–2570.
- Keitel A, Ince RAA, Gross J, Kayser C (2017) Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *Neuroimage* 147:32–42.
- Keitel A, Gross J, Kayser C (2018) Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biol* 16:e2004473.
- Ki JJ, Kelly SP, Parra LC (2016) Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *J Neurosci* 36:3092–3101.
- Kisler T, Reichel U, Schiel F (2017) Multilingual processing of speech via web services. *Comput Speech Lang* 45:326–347.
- Klein C, Liem F, Hänggi J, Elmer S, Jäncke L (2016) The “silent” imprint of musical training. *Hum Brain Mapp* 37:536–546.
- Klein ME, Zatorre RJ (2011) A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia* 49:878–887.
- Koelsch S, Fritz T, Schulze K, Alsop D, Schlaug G (2005) Adults and children processing music: an fMRI study. *Neuroimage* 25:1068–1076.
- Kraus N, Strait DL, Parbery-Clark A (2012) Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Ann N Y Acad Sci* 1252:100–107.
- Kreitewolf J, Mathias SR, Trapeau R, Obleser J, Schönwiesner M (2018) Perceptual grouping in the cocktail party: contributions of voice-feature continuity. *J Acoust Soc Am* 144:2178.
- Kubaneck J, Brunner P, Gunduz A, Poeppel D, Schalk G (2013) The tracking of speech envelope in the human cortex. *PLoS One* 8:e53398.
- Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8:194–198.
- Lankinen K, Saari J, Hlushchuk Y, Tikka P, Parkkonen L, Hari R, Koskinen M (2018) Consistency and similarity of MEG- and fMRI-signal time courses during movie viewing. *Neuroimage* 173:361–369.
- Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci* 31:2906–2915.
- Liu Z, Fukunaga M, de Zwart JA, Duyn JH (2010) Large-scale spontaneous fluctuations and correlations in brain electrical activity observed with magnetoencephalography. *Neuroimage* 51:102–111.
- Moore E, Schaefer RS, Bastin ME, Roberts N, Overy K (2014) Can musical training influence brain connectivity? Evidence from diffusion tensor MRI. *Brain Sci* 4:405–427.
- Mormann F, Lehnertz K, David P, Elger CE (2000) Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D* 144:358–369.
- Naci L, Sinai L, Owen AM (2017) Detecting and interpreting conscious experiences in behaviorally non-responsive patients. *Neuroimage* 145:304–313.
- Oechslin MS, Gschwind M, James CE (2018) Tracking training-related plasticity by combining fMRI and DTI: the right hemisphere ventral stream mediates musical syntax processing. *Cereb Cortex* 28:1209–1218.
- Palomar-García M-Á, Zatorre RJ, Ventura-Campos N, Buechekú E, Ávila C (2017) Modulation of functional connectivity in auditory-motor networks in musicians compared with nonmusicians. *Cereb Cortex* 27:2768–2778.
- Parbery-Clark A, Strait DL, Anderson S, Hittner E, Kraus N (2011) Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082.
- Parbery-Clark A, Tierney A, Strait DL, Kraus N (2012) Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219:111–119.
- Park H, Ince RA, Schyns PG, Thut G, 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.
- Patel AD (2014) Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hear Res* 308:98–108.
- Petrides M, Pandya DN (2009) Distinct parietal and temporal pathways to the homologues of Broca’s area in the monkey. *PLoS Biol* 7:e1000170.
- Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as asymmetric sampling in time. *Speech Commun* 41:245–255.

- Puschmann S, Baillet S, Zatorre RJ (2019a) Musicians at the cocktail party: neural substrates of musical training during selective listening in multi-speaker situations. *Cereb Cortex* 29:3253–3265.
- Puschmann S, Daeglau M, Stropahl M, Mirkovic B, Rosemann S, Thiel CM, Debener S (2019b) Hearing-impaired listeners show increased audiovisual benefit when listening to speech in noise. *Neuroimage* 196:261–268.
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724.
- Regev M, Simony E, Lee K, Tan KM, Chen J, Hasson U (2019) Propagation of information along the cortical hierarchy as a function of attention while reading and listening to stories. *Cereb Cortex* 29:4017–4034.
- Rosenthal G, Tanzer M, Simony E, Hasson U, Behrmann M, Avidan G (2017) Altered topology of neural circuits in congenital prosopagnosia. *Elife* 6:e25069.
- Sammler D, Grosbras MH, Anwender A, Bestelmeyer PE, Belin P (2015) Dorsal and ventral pathways for prosody. *Curr Biol* 25:3079–3085.
- Schiel F (1999) Automatic phonetic transcription of non-prompted speech. In: *Proceedings of the 14th International Congress of Phonetic Sciences* (Ohala JJ, ed), pp 607–610. Berkeley, CA: Department of Linguistics, University of California.
- Schlaug G (2015) Musicians and music making as a model for the study of brain plasticity. *Prog Brain Res* 217:37–55.
- Schön D, Magne C, Besson M (2004) The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* 41:341–349.
- Seghier ML (2008) Laterality index in functional MRI: methodological issues. *Magn Reson Imaging* 26:594–601.
- Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, Hasson U (2016) Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* 7:12141.
- Strait DL, Kraus N (2011) Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front Psychol* 2:113.
- Strait DL, O'Connell S, Parbery-Clark A, Kraus N (2014) Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cereb Cortex* 24:2512–2521.
- Swaminathan S, Schellenberg EG (2017) Musical competence and phoneme perception in a foreign language. *Psychon Bull Rev* 24:1929–1934.
- Tadel F, Baillet S, Moshier JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci* 2011:879716.
- Varnet L, Wang T, Peter C, Meunier F, Hoen M (2015) How musical expertise shapes speech perception: evidence from auditory classification images. *Sci Rep* 5:14489.
- Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N (2011) What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. *Neuroimage* 54:577–593.
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10:420–422.
- Xu J, Kemeny S, Park G, Frattali C, Braun A (2005) Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage* 25:1002–1015.
- Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6:37–46.
- Zatorre RJ, Chen JL, Penhune VB (2007) When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci* 8:547–558.
- Zhao TC, Kuhl PK (2016) Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proc Natl Acad Sci U S A* 113:5212–5217.