

# The Topographic Representation of Time and its Link With Temporal Context and Perception

**Shrikanth Kulashekhar**

International School for Advanced Studies (SISSA)

**Sarah Maass**

University of Groningen

**Hedderik van Rijn**

University of Groningen

**Domenica Bueti** (✉ [domenica.bueti@sissa.it](mailto:domenica.bueti@sissa.it))

International School for Advanced Studies (SISSA) <https://orcid.org/0000-0001-6660-7241>

---

## Article

**Keywords:** topography, auditory chronomaps, temporal context, motor area, chronotopic maps

**Posted Date:** October 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-753370/v2>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Neuronal tuning and topography are mechanisms widely used in the brain to represent sensory information and also abstract features like time. In humans, temporal topography has been shown in a wide circuit of brain regions. However, it is unclear whether chronotopic maps are specific to vision, whether they map time in an absolute or relative fashion, to what extent they reflect objective or subjective time and whether they are influenced by temporal context. Here we asked human participants to reproduce the durations of sounds in two, partially overlapping, temporal contexts while we record high-spatial resolution fMRI. Both model-based and data driven analyses show the presence of auditory chronomaps in the auditory parabelt, intraparietal sulcus, and in supplementary motor area. Most importantly, when the same physical duration is presented in different temporal contexts, and thus perceived differently, different neuronal units respond to it. Those units are also spatially shifted according to the relative position of the perceived duration within each context. Finally, the pattern of activity is more similar within rather than across contexts suggesting their pivotal role in shaping the maps. These results highlight two important properties of chronomaps: their flexibility of representation and their dependency on the context.

## Introduction

The speed at which we move while dancing depends not only on our capacity of keeping the musical beat but also on the speed at which our dancing partner moves. Keeping the musical tempo while moving in sync with the dancing partner requires the rapid processing of multiple durations whose perception is prone to biases depending on the temporal features of the environment (i.e., how the current tempo relates to earlier perceived tempos). How the brain encodes and reads out the rapid succession of different durations and how the resulting perception is influenced by the temporal features of the environment is far from clear [1],[2].

Recently, electrophysiological work in animals and neuroimaging studies in humans have shown the existence in cortical and subcortical brain areas of a form of duration tuning, that is, neuronal units selectively responsive to different durations [3],[4],[5]. In humans these units are also topographically organized on the cortical surface to form chronomaps [6],[7]. Chronotopic maps associated with visual temporal discrimination tasks have been observed in the Supplementary Motor Area (SMA [6],[8]), whereas chronotopic maps linked to stimulus variation of both duration and temporal frequency have been reported in a wide network of cortical areas, that is, from occipital to parietal to prefrontal regions [7].

Although these studies did show the existence of a topographical representation of time in the human brain, a number of questions are still open. First, it is unclear whether chronomaps are specific to vision and to discrimination tasks or whether they are “amodal” and exist independently from the sensory modality and the task at hand. Second, is time mapped in relative or absolute terms in chronomaps? Is the physical duration or the relative position of the duration within a distribution that is reflected in the

chronomaps? Third, is the activity in chronomaps modulated by how a duration is perceived according to changes in the temporal features of the environment (i.e., temporal context)? In parallel to these questions, we will consider in what way the duration representations equate or differ from other topographic representations in the brain.

To address these questions, we used a temporal reproduction task of sounds of different durations (five durations ranging from 0.32 to 1.1 s) and asked participants to reproduce them in two partially overlapping temporal contexts. In the *short context*, the durations ranged from 0.32 to 0.65 s and in the *long context* they ranged from 0.65 to 1.1 s, with the 0.65 s duration appearing in both contexts.

By manipulating the temporal context, the perception of the 0.65 s duration (i.e., the duration shared between contexts) will be biased towards the mean value of the duration distribution of each context. It will be perceived as shorter when presented in the short context, and as longer in the long context [9],[10]. This effect, which is called “central tendency” or “regression to the mean” has been interpreted within a Bayesian framework as an optimization mechanism that takes into account the knowledge of the duration distribution at hand to provide an accurate duration estimation of the sample stimulus [11],[10], [12]. Although this effect of regression to the mean has been extensively documented at a behavioral level, and recent EEG studies have focused on traces of the Bayesian integration [11], whether and where there is a neural signature of the subjectively modified duration is unclear. Specifically, it is unclear if and how the representation of the same physical duration changes when its perception changes as a function of the temporal context, and how the brain responds to durations that are grouped according to it.

In summary, our experimental design enabled us to answer: (a) whether chronomaps extend to durations perceived in the auditory domain, (b) whether they map time in absolute or relative terms as the distinct temporal contexts will allow for testing whether the representation of a physical duration changes when it is perceived differently, and (c) whether the representation of different durations in a map is influenced by the presence of distinct temporal contexts.

Based on the above rationale, we identified four possible predictions about the topographic representation of time and its interaction with perception and temporal context (see Fig. 1A):

1. If chronomaps represent time in a veridical, absolute fashion, the context should not affect duration representation. We should thus observe a single map with different voxels active for each of the *physical duration* in the two contexts, with the 0.65 s duration eliciting activation in the same cluster of voxels.
2. If chronomaps are quantitative representations of the subjective experience of time, we expect different voxels to be active for the different *perceived durations*. Due to the central tendency effect elicited by the two contexts, we would expect different voxels to be active for 0.65 s duration, resulting in a single six-duration map.
3. If time is represented in a categorical, *relative* fashion (e.g., voxels representing “shortest”, “intermediate”, “longest” durations) that is irrespective of the context, we expect a single map for

both contexts, where the same voxels are active for durations that have the same relative position within the context (i.e., shortest, intermediate, longest).

4. If time is represented in a relative fashion but its representation interacts with the context, we expect two maps, one for each context, with a varying degree of overlap. The presentation of 0.65 s should then elicit the activation of different voxels whose location corresponds to the position of this duration within the appropriate context.

## Materials And Methods

### Participants

Fourteen healthy, right-handed volunteers (mean age  $23 \pm 3$  years, mean  $\pm$  standard deviation, seven females) participated in the study. All volunteers gave written informed consent to participate in this study, the procedures of which were approved by the International School for Advanced Studies (SISSA) ethics committee (protocol number 1899/II-16) in accordance with the Declaration of Helsinki.

### Stimuli and Procedure

We used an auditory temporal reproduction task, in which subjects were asked to reproduce, by pressing, holding down and releasing a response key, the duration of a sound (a pure tone 1000 Hz in pitch) delivered via headphones. The beginning of a trial was indicated by a visual cue, an 'X' ( $2^\circ$  of visual angle), presented on the screen placed at the posterior end of the MRI bore and lasting for 0.2s (Fig. 1A). After a brief post-cue period (1.3–2.3 s), a single pure tone was played via headphones for a variable duration (ranging from 0.32 to 1.1 s). After an interval ranging from 2 to 4 s, a burst of white noise presented for 0.1s instructed the subjects to reproduce the previously heard sound by pressing and holding down a response key. The subjects did not receive any feedback on their performance after their response. After the response, the next trial started following an inter-trial interval ranging from 0.3 to 0.5s. Occasionally, the subjects were randomly presented with catch trials in which only the pure tone was presented. In the catch trials the subjects did not reproduce the duration.

Subjects were tested separately in two temporal contexts. In the short context, the sound's duration was either 0.32s, 0.46s or 0.65 s, in the long context it was 0.65 s, 0.85 s or 1.1 s. 0.65 s was presented in both contexts and it was the shortest duration in the long context and the longest duration in the short one. Every fMRI run consisted of 54 experimental trials and 9 catch trials, with 18 experimental trials and 3 catch trials for each duration; all trial types and durations were presented randomly. We collected 3 fMRI runs for each temporal context. The 3 runs of each context were always presented in sequence, whereas the presentation order of short and long context was counterbalanced across subjects. A total of 378 trials were collected for every subject, with 189 trials for each context and 63 trials for each of the 6 durations. The experimental paradigm was designed and presented using the Psychophysics toolbox [33] in Matlab (The Mathworks, Inc.).

### Behavioral Data Analysis

For each participant and each duration of the two temporal contexts, we took as a measure of accuracy the reproduced duration, which was the time between response key press and response key release. To check for significant differences in the reproduced duration between the two contexts, the individual reproduced durations were entered in a repeated measures ANOVA with two contexts (short and long) and three durations (short, intermediate and long duration) as factors. As post-hoc tests we used paired t-tests in which the alpha level was set to 0.05.

## **MRI Acquisition**

Blood oxygenation level-dependent (BOLD) functional imaging was performed using an actively shielded, head-only 7T MRI scanner (Siemens, Germany), equipped with a head gradient-insert (AC84, 80 mT/m max gradient strength; 350 mT/m/s slew rate) and 32-channel receive coil with a tight transmit sleeve (Nova Medical, Massachusetts, USA). The ultra-high magnetic field system allowed us to have voxels with smaller size compared to lower field MRI thus increasing the spatial resolution of the functional data. Moreover, in 7T systems the signal strength of venous blood is reduced due to a shortened relaxation time, restricting activation signals to cortical grey matter which results in a better signal-to-noise ratio[34]. Time-course series of volumes were acquired for each run using the multiband sequence. The spatial resolution was 1.5 mm isotropic, the volume acquisition time (TR) was 1368 ms, the flip angle was 60 degrees, the echo time (TE) 23 ms and the bandwidth 1903 Hz/Px. The matrix size was 146 x 146 x 75, resulting in a field of view of 219 (AP) x 219 (RL) x 112.5 (FH) mm. An undersampling factor 0 and CAIPIRINHA shift 3 were used. Slices were oriented transversally with the phase-encoding direction anterior-posterior. 146x42x75 reference lines were acquired for the GRAPPA reconstruction.

High-resolution whole-brain MR images were also obtained using the MP2RAGE pulse sequence optimized for 7T (voxel size = 0.60 x 0.60 x 0.59 mm, matrix size 320 x 320 x 256,  $T_1/T_2 = 750/2350$ ms,  $\alpha_1/\alpha_2 = 4/5$  degrees,  $TR_{MP2RAGE}/TR/TE = 5500/6000/4.94$  ms).

## **fMRI Preprocessing**

Functional imaging data were preprocessed using the Statistical Parametric Mapping (SPM12 v. 7219, Wellcome Department of Imaging Neuroscience, University College London) toolbox in MATLAB. In each individual subject the EPI volumes acquired in the different runs were first realigned. The runs were first realigned to each other, by aligning the first scan from each session to the first scan of the first session. Then the images within each session were aligned to the first image of the session. The realigned images were then co-registered to the T1-weighted image acquired in the same session. The subject's images in native space realigned and co-registered to the T1-weighted image were next smoothed with a 2 mm full-width at half-maximum Gaussian kernel.

## **GLM analysis**

The fMRI time series were analyzed at individual subject level using a univariate GLM approach. The events of interest in the GLM analysis included the offsets of the three durations in the two contexts and

the onset of the response (i.e., onset of the keypress). We also modelled the visual cue onset signaling the beginning of each trial and the six motion correction parameters as effects of no interest. The duration of all events was set to zero. The individual GLM included the six fMRI runs, three for each context, and each run had 13 regressors (7 of interest and 6 of no interest). All events were convolved with the canonical hemodynamic response function (HRF). The fMRI time series data were high-pass filtered (cutoff frequency = 0.0083 Hz). Correction for non-sphericity [35] was used to account for possible differences in error variance across conditions and any non-independent error terms for the repeated measures. To identify the brain areas exclusively active at the offset of the encoded sound (and not during reproduction) independently from the different durations and the two contexts we contrasted duration and response (duration offset - response onset, resulting in one t-contrast for each subject) and we averaged across durations and contexts. To identify the presence of auditory chronomaps in each temporal context i.e., voxels exclusively active at the offset of the sound, but not during the reproduction, and maximally activated by each specific duration we used as contrast of interest sound offset - response onset for each sound and context (resulting in 6 t-contrasts for subject). In all t-contrasts,  $p_{FWE} < 0.05$  corrected for multiple comparisons across the whole brain.

*Winner-take-all.* To appreciate the existence of chronomaps in each temporal context, the three t-maps, obtained at single subject level and for each context were then used to classify the voxels according to their preference to one of the 6 different durations (three durations in the two contexts). Voxels were classified according to a “winner take all” rule (WTA), for example voxels with the greatest t value (threshold was set to  $t > 3.13$ ), for the shortest duration range in the short context (0.32 s) were classified as responsive to that duration range and labeled with number 1. We created 6 different labels corresponding to each duration in the two contexts i.e., 0.32s, 0.46s, 0.65 s (SC), 0.65 s (LC), 0.85 s and 1.1 s. For WTA we used only the clusters of voxels that were significant at  $p < 0.05$  cluster-level corrected for multiple comparisons across the whole brain.

*Second-level GLM analysis.* To make sure that the activations observed at individual level also emerged as a statistically robust group level result we also ran a GLM 2-level analysis. For each subject, we computed a GLM analysis on MNI normalized images. The GLM analysis was identical to that described earlier. The MNI normalization was computed using the SPM 12 pipeline. The Realigned and normalized images were smoothed with a 4 mm full-width at half-maximum Gaussian kernel.

For each subject we estimated 12 t-contrasts, one for each duration and each response in the two temporal contexts, averaging across the 3 runs of each context. The resulted contrast images (t-statistics maps) were then used in a second-level ANOVA where we contrasted all durations against all responses (6 durations – 6 responses). The statistical threshold was set to  $p < 0.05$  FWE cluster-level corrected for multiple comparisons across the entire brain volume (cluster size estimated at a voxel level threshold  $p$ -uncorrected = 0.001). Correction for non-sphericity (Friston et al., 2002) was used to account for possible differences in error variance across conditions and any non-independent error terms for the repeated-measures.

## Anatomical image processing

The high-resolution MP2RAGE images were analyzed using Freesurfer software [36] (<http://surfer.nmr.mgh.harvard.edu/>). Freesurfer's automatic pipeline performs the volumetric segmentation of the MRI data, the surface reconstruction of inflated surfaces, the flattening of cortical regions of interests, the cortical parcellation, and the neuroanatomical labelling with the Freesurfer/Destrieux atlas [37].

*Morphing using Freesurfer.* Visualizations and computations requiring moving surface data from different subjects into a common surface were performed using the Freesurfer operation "*mri\_surf2surf*". The source surface was the surface closest to the mean surface estimated across subjects. For example, when morphing data of the SMA Freesurfer label from different subjects into one destination subject space, the SMA label area was first estimated for all subjects. The destination subject space was the subject with the SMA label closest to the mean of the SMA labels across subjects. This method of morphing ensures the best transformation of data from multiple sources to a single destination space.

## Surface-based quantification of chromomaps spatial progressions

Chromomaps were visualized and their metrics estimated on inflated and flattened cortical surfaces.

The areas where we explored the existence of chromomaps, which were significantly active at the offset of all sounds and contexts in all subjects (see S-Fig1) were called Region of Interest (ROI). These were SMA, IPS and parabelt areas.

Chromomaps were identified in the left and the right hemisphere of each individual subject using the SPM t-maps resulting from the WTA method. These volumetric maps were projected onto the cortical surface of each individual brain following the Freesurfer pipeline (with a projection fraction set to 0.5). Individual chromomaps for short and long context separately were identified in SMA, IPS and parabelt areas of both hemispheres. Maps in all subjects, ROIs and hemispheres were visually identified when there was a clear spatial progression of duration preferences from short to long durations. Maps' borders were manually drawn at the edge of the clusters of vertices (vertices are voxels projected into the cortical surface) which preferred the longest and the shortest duration of the range.

The spatial progression was quantified on flattened surfaces as the normalized distance (*nD*) of each duration selective vertex from the shortest edge for the map. The normalized distance was defined as:

$$nD = \frac{\text{Vertex Distance from Shortest Edge}}{\text{Distance between short \& long edges}}$$

The distance was computed for each vertex in the cluster and then averaged across vertices of the same cluster. The average vertex distance was then estimated for each duration selective cluster of vertices in the two contexts in all subjects and tested for statistical significance using a t-test. In each individual subject, a slope of the spatial progression of duration selective vertices was also computed.

To quantify the spatial shift of the duration selective clusters active for the duration shared between the contexts, for each individual subject and for each context separately we computed the distance of the 0.65 s cluster (averaging the distances of all vertices within the cluster) from the shortest edge of the map and we then compare it across subjects using a t-test. To check for the presence of a correlation between spatial shifts in the cortical representation of 0.65 s and reproduction of this duration across contexts, we plotted the difference in the reproduced duration of 0.65 s in the two contexts (SC - LC) against the absolute difference in distance of the 0.65 s duration selective clusters in the two contexts.

*SMA chronomap.* In SMA, whose location was double checked with the Freesurfer “BA6” label of both the hemispheres, we defined chronomap’s orientation as the spatial progression of clusters of vertices showing duration preferences. A flat surface of the “BA6” label was created for each hemisphere and subject.

*IPS chronomap.* Chronomaps in the IPS were identified using the Freesurfer’s IPS label (“S\_intrapariet\_and\_P\_trans” label) for each subject. For each subject and hemisphere, a flat surface representation of the IPS label was then created to compute the map’s attributes. To have a more data-driven approach in determining the chronomap progressions and orientations we developed an octagonal search method to determine the best map orientation in the IPS. An octagonal search grid was assumed over the IPS area. The eight sides of the octagon then served as the borders or edges for possible test orientations, resulting in four pairs of shortest-to-longest borders. The primary orientation was assumed to be orthogonal to the postcentral gyrus (poCG). The orientations were defined as relative to this primary orientation axis, with the remaining three axes of the octagon at 45°, 90° and 135°. For each context, the average vertex normalized distances (nD) were computed across the durations and octagonal test orientations. For every test orientation, a slope was computed from the nDs of the different durations in the two contexts. The slope reflected how well the duration clusters were topographically organized in that orientation. The winning map orientation for a given subject and hemisphere was the orientation resulting in the steepest slope. This method of using a common anatomical reference makes the resulting map orientations comparable across the subjects.

*Parabelt chronomap.* The auditory parabelt was defined as the ROI including the following Freesurfer labels: “G\_temp\_sup-Lateral”, “S\_temporal\_transverse”, “Lat\_Fis-post”, and “G\_temp\_sup-Plan\_tempo”. A similar octagonal search grid method used with the IPS chronomaps was applied to the auditory parabelt maps for each hemisphere and subject. The primary orientation here was assumed to be orthogonal to Heschl’s gyrus (HG). For every test orientation, a slope was computed from the average nDs of the different durations in the two contexts. The winning orientation for a given subject and hemisphere was the orientation resulting in the steepest slope.

*Overlap between the two context maps.* To visualize the extent of segregation and overlap between the short and long contexts, the two context maps in each ROI (i.e., SMA, IPS and parabelt) were visualized together on individual cortical surfaces. When the two context maps were visualized at group level, for each subject, only the hemispheres with the winning orientation were overlaid on the common surface

space. The dominant orientation in SMA was the anterior-to-posterior (N left hemispheres = 19, N right hemispheres = 12), in IPS it was the orientation orthogonal to the poCG (N left hemispheres = 8, N right hemispheres = 6) and in the parabelt areas it was the orientation parallel to the HG (N left hemispheres = 4, N right hemispheres = 3).

Moreover, to quantify the amount of overlap between the maps in the two contexts, we computed the difference of the shortest and longest edges of the maps in two contexts i.e., SC-LC. This difference was computed for each ROI in each individual flat surface using only the maps where both short and long contexts had the same dominant orientation. We first estimated the distance of each map border (i.e., shortest and longest edge) from an anatomical landmark, which was pCG, poCG, HG for SMA, IPS and parabelt, respectively. We then subtracted this distance value of the long context from the same distance value of the short context (SC-LC).

### Duration tuning analysis

We checked the response properties of duration selective clusters of voxels by also looking at the BOLD response in those clusters to preferred and non-preferred durations within and across contexts. In each subject, ROI and context to avoid circularity, the duration selective clusters of voxels were identified in one run and the hemodynamic response of those clusters was extracted from the remaining two runs (in all possible combinations). The duration selective clusters from a single run were identified using the GLM analysis and WTA approach, as described earlier (see GLM and winner-take-all analysis). For each cluster of duration-selective voxels the normalized hemodynamic response was estimated as:

$$BOLD(t) = \frac{\sum_{i=1}^{Nruns} \sum_{v=1}^{Nvoxels} \frac{x(t) - MB}{MB}}{Nruns * std \left( \sum_{v=1}^{Nvoxels} \frac{x(t) - MB}{MB} \right)}$$

Where,  $x(t)$  is the signal in each voxel and MB is the baseline that was obtained averaging the signal of  $t$  for each run. Normalization was performed by subtracting the signal in each voxel from a baseline value and dividing it by the baseline. The BOLD response was aligned to the second volume (i.e., a TR) after the duration offset (see also <sup>6</sup>). Within a single subject, we first averaged the BOLD signal across the voxels of a cluster and then across the fMRI runs.

### Multivariate Pattern Recognition Analysis (MVPA)

The multivariate pattern analysis (MVPA) was performed using the CosmoMVPA toolbox[38] in MATLAB (Matlab Inc.). For the MVPA analysis the fMRI time series were reanalyzed as before (see GLM analysis) but in the fMRI preprocessing the realigned and co-registered images were unsmoothed. The GLM analysis, as specified above, included the six runs from the two contexts. Each run included 7 events of interest and 6 events of no interest. The modelled events included the visual cue onset, the three sounds offset, the three response onsets and the six motion correction parameters. The beta values associated with the offset of the six durations (three for each context) were then used for the MVPA analysis.

*Predicting the different durations in the two contexts.* In the first place we used MVPA to check whether from the activity of SMA, IPS and parabelt areas we could predict the three different durations in the two contexts. To this purpose we used a leave-one-run-out cross-validation approach. For each subject and in each ROI, a support vector machine (SVM) classifier (LIBSVM [39]) was trained to classify the pattern of activity associated with the six durations from two runs (one for each context). The classifier was then tested using the activity pattern from the left-out run. This classification routine was iteratively performed until every run was left out once (3 iterations). The overall classification accuracy was then computed by averaging the classification accuracy from all iterations. The classification accuracies resulting from this analysis were visualized as confusion matrices (chance level is  $\chi=0.17$ ). To test if the cross-validation results were specific to the ROIs associated with the timing task, we performed the same analysis on two additional task-unrelated ROIs. The two control ROIs were the occipital pole (OP) and the orbitofrontal cortex (OC). OC was defined using the “G\_orbital” and “S\_orbital-H\_Shaped Freesurfer” labels. While the OP was defined using the Freesurfer “Pole\_occipital” label. To compare the prediction accuracy of the different ROIs we used 6 durations by 6 ROIs, repeated measures ANOVA and we used paired t-tests as post-hoc tests. Alpha level was set to  $p=0.05$ .

*Predicting the contexts.* With MVPA we also tested if the pattern of activity in SMA, IPS and parabelt and the two control ROIs could predict the two contexts independently from the different durations. As before, we used a cross-validation approach. Here the SVM classifier was trained to classify the pattern of activity associated with the two contexts from two runs and then tested using the activity pattern from the left out run. A classification accuracy at chance level was equal to  $\frac{1}{2}=0.5$ .

*Double-checking the duration preferences.* We performed a complementary MVPA analysis to establish if the duration selective clusters identified with the winner-take-all approach were indeed selective to the assigned duration [40]. This analysis was performed using a cross-validation approach as described before. Here, instead of using the whole ROIs, the cross-validation was carried out separately for each duration selective cluster of SMA, IPS and parabelt areas. To check the decoding accuracy of the different duration selective clusters of voxels, this searchlight analysis was conducted with a search area of 1 voxel that moved across the whole duration selective cluster. For each of the voxels, the cross-validation analysis was carried out using a leave-one-out approach. An SVM was trained to classify the pattern of activity associated with the six durations from two runs and then tested using the response pattern from the left out run. For all subjects and all ROIs, we estimated the classification accuracy of each voxel in all duration selective clusters.

## **Representational similarity analysis**

To measure how the brain activity differed between the six different durations in the two contexts, we analyzed the fMRI time series using a multivariate representational similarity analysis (RSA) [41]. With RSA for each ROI i.e., SMA, IPS, parabelt, we correlated the beta values associated with the offset of each duration with the beta values associated with every other offset duration within and across contexts. To perform this correlation, in each individual subject we averaged the betas corresponding to each duration

offset across the three runs of the same context. The correlation was measured with the Spearman correlation coefficient. The resulting correlation coefficients were entered into a representation dissimilarity matrix (RDM) where each entry was created by subtracting the correlation coefficient by 1 and averaging this correlation coefficient across subjects. This value reflects how dissimilar on average each duration representation is from the others.

## Results

To test the existence of auditory chonomaps and the relationship of this topographic representation with temporal context and perception, we asked 14 healthy volunteers to perform a temporal reproduction task of pure tones of different durations (see Fig. 1B and Material and Methods section for more details) while we recorded high spatial resolution fMRI images with a 7 Tesla MRI scanner.

In different fMRI runs, volunteers were presented with sounds of different durations (i.e., three runs for each temporal context). In the short context the sounds ranged from 0.32 to 0.65 s and in the long context they ranged from 0.65 to 1.1 s. The presentation of the sound (a pure tone) was followed by a reproduction phase in which volunteers, after being cued with a burst of white noise (0.1 s), had to press and hold a response key down for a period of time matching the duration of the previously heard sound.

Behaviorally, the results (see Fig. 1C) are in line with the expected regression to the mean within each context (3 durations by 2 contexts repeated measures ANOVA; main effect of context  $F(1,2) = 63.3$   $p < 0.001$   $\eta^2 = 1.01$ ). Overall the mean reproduction in the shortest context was shorter ( $0.76 \text{ s} \pm 0.24 \text{ s}$ ; mean  $\pm$  standard deviation) compared to the longest context ( $1.4 \text{ s} \pm 0.4 \text{ s}$ ; mean  $\pm$  standard deviation). In both temporal contexts the shortest duration was overestimated and the longest duration underestimated (main effect of duration  $F(1,2) = 11.45$ ,  $p < 0.001$ ,  $\eta^2 = 5.59$ ), whereas the reproduction of 0.65 s was significantly different in the two contexts (paired t-test  $t(13) = -4.92$ ,  $p < 0.001$ ); 0.89s (SD: .23) in the short context and 1.1 s (SD: .31) in the long context, confirming that the same physical duration (0.65 s) was perceived differently in the different contexts.

At the neural level we performed a General Linear Model (GLM) analysis for each subject individually with the offset of the three sounds in the two contexts and the onset of their reproduction as events of interest. All regressors were convolved with the canonical hemodynamic response function (for the modelling of the other events, see Material and Methods). In each individual subject we first looked for brain areas exclusively active at the offset of the encoded sound (and not during reproduction) independently from the different durations and the two contexts (i.e., the contrast of interest for each context was: 3 durations – 3 reproductions, the contrast was  $p_{FWE} < 0.05$  corrected for multiple comparisons across the whole brain). As in previous work [6], [8], we modeled the event offset because it was the moment at which the duration of the sound became available to participants. We contrasted the sound offset with the response onset to remove activations purely correlated with the motor response and with the mere auditory processing of the burst of white noise signaling the beginning of the reproduction phase. The result of this contrast, which reflects then the appreciation of the auditory durations, revealed in the

majority of subjects the activation of the Supplementary Motor Area (SMA), of the inferior parietal lobule (specifically the intraparietal sulcus, IPS) and, less consistently across individuals, of the auditory parabelt areas. Since our main interest was the identification of chronomaps and their change with temporal context and perception, all our fMRI analyses were conducted at single subject level in the native space. However, in S-Fig. 1 and S-Table 1 we also show the results of a group level analysis we ran to make sure that the activations observed at individual level also emerged as a statistically robust group level result.

We then focused on individual SMA, IPS and parabelt areas, whose location was double checked with available atlases (see Materials and Methods for more details), to identify the presence of auditory chronomaps in each temporal context, that is, voxels exclusively and maximally active at the offset of the sound but not during the reproduction onset nor during the auditory processing of the cue signaling the reproduction phase, (i.e., the contrast of interest for each sound and context was sound offset - response onset, the contrast was  $p_{\text{FWE}} < 0.05$  corrected for multiple comparisons across the whole brain). The two temporal contexts were considered separately because they were tested separately (i.e., participants never experienced the 6 durations together, see Materials and Methods for more details).

Figure 2A shows the bilateral SMA for the short (leftwards panel) and the long (rightwards panel) context, and highlights a number of individual maps (for all individual maps, see S-Fig. 2 and S-Fig. 3). Figure 2A shows color-coded the cluster of vertices (voxels projected onto the cortical surface) classified as maximally responsive to each of the three sound durations, based on a t-statistic winner-takes-all procedure. The color scale ranges from red, corresponding to vertices responsive to the shortest duration (0.32 and 0.65 of the short and the long context respectively), to blue, the vertices maximally responsive to the longest duration (0.65 s and 1.1 s of the short and the long context respectively). The maps were characterized by the presence of a spatial transition in duration preferences, that is, from shortest to intermediate to longest duration in a given context. The borders of the maps were drawn for each individual subject on the basis of this transition for each hemisphere and context (all analyses that are relative to the maps' identification and the computation of distances were done at surface level).

Figure 2B shows the distance of duration selective clusters for each context (i.e., the average distance of all vertices in a cluster, see Material and Methods for more details) from the shortest border in each individual map (i.e., the border closest to the shortest duration preference, the dashed border of Fig. 2A) and also as an average (continuous line). In both contexts and hemispheres, there was a clear spatial progression from the shortest to the longest edge border (all t-tests  $p < 0.01$ ). This progression was in the anterior to posterior direction with vertices preferring the longest duration in the context and closer to the precentral gyrus compared to those preferring the shortest duration in the majority of the SMA maps (67%) (see Fig. 2A). Although the anterior to posterior orientation was most prominent, other map's orientations were observed in a minority of maps (see Supplementary Fig. 15A). It is worth emphasizing here that compared to Protopapa et al. 2019 [6], we have allowed more flexibility in the identification of map's borders following the methodology described in Harvey et al. 2020 [7].

Once we assessed the existence of auditory chronomaps in each temporal context, we moved to explore the representation of 0.65 s duration in the two maps (Fig. 3A, B). Where is the same physical duration represented in the two contexts? Fig. 3A shows the vertices responding to 0.65 s, color-coded according to the context for a few individual subjects (top row) and for the group (bottom of panel A). In blue are the vertices that are maximally responsive to 0.65 s in the short context (i.e., longest duration of the distribution) and in red those responsive to 0.65 s in the long context (i.e., shortest duration of the distribution). In green are the vertices that keep the same preference across the contexts. The figure shows that the presentation of 0.65 s elicits the activation of different vertices, according to the relative position of this duration within each context, with more anterior clusters of vertices preferring 0.65 s in the long context and more posterior vertices responding to 0.65 s in the short context. This spatial shift was measured in each individual map and hemisphere as the distance of vertices preferring 0.65 s in each context from the shortest border of the appropriate map (Fig. 3B, see S-Fig. 4 to see all individual maps). The vertices preferring 0.65 s in the short context map were located more posteriorly compared to the vertices preferring 0.65 s in the long context map. We then checked whether the spatial shift of the clusters responding to 0.65 s in the two contexts correlated with perceptual differences (i.e., with the reproduction of 0.65 s in the two contexts). As shown in Fig. 3C, at least in the left hemisphere, this correlation was significant ( $r = -0.9$ ;  $p = 0.004$ ; Pearson correlation). The greater the difference in reproduction (i.e., the more negative values) the greater the spatial difference between the clusters.

**(B)** Shows the normalized distance of the 0.65 s duration selective clusters in SC and LC. In red the clusters of the LC, in blue those in the SC, for simplicity, the data of the right and left hemispheres are considered together ( $N(\text{SC}) = 15$  and  $N(\text{LC}) = 21$ ). The circles are individual subjects/hemispheres, the black points their means. Each circle is the average distance of all vertices in a 0.65 s cluster. The normalized distances were computed in each subject, hemisphere and context separately as the difference (normalized to the map's size) from the shortest edge of the map.  $***p < 0.001$  **(C)** Shows the correlation between the difference in reproduction time (time from key-press to key release) of 0.65 s (y axis) and the normalized distance of the 0.65 s clusters in short and long contexts. Data points are individual subjects and the dashed line is the best fit line. In the leftward panel is the left hemisphere, in the rightward is the right hemisphere. For the correlation, we considered for each hemisphere only the subjects that had significant activation of the 0.65 s cluster in the two contexts ( $N = 7$ ).

At this point, to rule out the possibility that the spatial shift of 0.65 s representation in the two contexts was random, we decided to check the consistency of the maps' spatial progression within and across contexts. In each subject we checked the correlation of the maps spatial progression (i.e., the slopes resulting from the computation of the distances of each duration selective cluster from the shortest edge of the map) between the different runs of the same and of different contexts (see S-Fig. 5). As expected, the spatial progression of the maps was highly correlated between the runs of the same context (for short context  $r_{\text{run1,run2}} = 0.5$ ,  $p = 0.11$ ;  $r_{\text{run2,run3}} = 0.59$ ,  $p = 0.04$ ;  $r_{\text{run1,run3}} = 0.71$ ,  $p = 0.007$ ; long context  $r_{\text{run1,run2}} = 0.46$ ,  $p = 0.04$ ;  $r_{\text{run2,run3}} = 0.6$ ,  $p = 0.005$ ;  $r_{\text{run1,run3}} = 0.49$ ,  $p = 0.05$ ) where no change was expected. Across

the contexts this correlation was much lower and less consistent ( $r$  ranging from 0.5  $p = 0.07$  to -0.20  $p = 0.45$  see S-Fig. 5 for more details).

The fact that different vertices are active for the same physical duration when perceived differently and that those vertices are spatially shifted according to the relative position of this duration within each context, suggests that time is represented in a relative fashion within the maps. However, it remains unclear the extent of this relative representation of time i.e., how much overlap does exist between the maps in the two contexts? And what is the role played by the context in shaping the maps? If the context does not play any role, the maps in the two contexts should be totally overlapping, otherwise they should show a certain degree of separation (see Fig. 1A).

Figure 4AB shows the overlap between the maps of the two contexts (see S-Fig. 6 to see the overlap between contexts in individual subjects). In orange color are the clusters active in the short context (SC) and in pale blue those active in the long context (LC), in yellow are the overlapping vertices. The two maps are neither spatially segregated nor totally overlapping. When we looked at the differences between short and long contexts borders (Fig. 4B), we see that in the majority of the subjects in which the map orientation was in the anterior to posterior direction in the two contexts ( $N = 8$ ), the posterior borders overlapped across contexts (in 6 out of 8 subjects, the difference was close to 0 in the  $y$  axis) but for the anterior borders the situation was more mixed, since in half of the subjects the anterior borders was more anterior in short compared to long context (positive values in  $x$  axis) and in the other half it was the reverse.

To better assess the overlap of the maps between contexts we looked at the hemodynamic response of duration selective voxels for preferred and non-preferred duration within, but most importantly, across contexts. For each individual subject to avoid circularity, the selection of the duration selective clusters was based on a single run and the hemodynamic response extraction was computed on the remaining runs in all possible combinations. Figure 4C shows for the average of the subjects, consistent duration preferences across the runs (i.e., different runs are the different symbols in the plot), with a gaussian-like type of response profile within but not across the contexts. As expected, the hemodynamic response was greater for the preferred duration and slowly decayed with distance from it. If the flexibility of duration representation in the maps was absolute, we should have seen changes in duration preferences across the contexts. Specifically, we expected the same cluster of voxels to peak for durations that, in the two contexts, were in the same relative position of the distribution. However, this is not the case; voxels did not change their duration preferences across contexts (i.e., the hemodynamic response only peaked for the appropriate duration within the context). This result was also corroborated by the observation that only 18.6% of the significantly active voxels were shared between the two contexts and only 17% of them were active for the same relative position within the context.

Overall the SMA results showed the presence of auditory chronomaps in the two contexts. The spatial progression was, as expected according to previous work [6], in the anterior to posterior direction from shorter to longer durations. Within each context, duration selective voxels show a gaussian-like type of

tuning, where response was enhanced for preferred duration and slowly decayed with distance from it. This result shows that chronomaps in SMA are sensory modality and task-independent. Different vertices respond to the same physical duration when it is perceived differently in the two temporal contexts. Those clusters of vertices are also spatially shifted according to the relative position with each context (i.e., anterior for the shortest duration and posterior for longest duration of the context). The spatial shift of the clusters in the left hemisphere correlates with the perceived difference of the same duration in the two contexts. This result seems to suggest that time is represented in a relative fashion within the maps. However, the observation that there is no change in tuning across contexts and that only few significantly active voxels are shared between them, suggests a pivotal role of the temporal context in shaping the activity within each map. Time representation is thus flexible but this flexibility is not absolute.

We then move to explore the existence of auditory chronomaps in auditory parabelt areas and IPS (for the details of the ROIs selection see Material and Methods). As for SMA, also in these areas, maps were identified by the presence of spatial transitions in duration preferences (i.e., from shortest to longest duration) and maps' borders were drawn in each individual subject, each hemisphere and context. Figure 5 shows the presence of chronomaps in IPS (Fig. 5A) and parabelt (Fig. 5B) in a few individuals (for all individual maps see S-Figs. 7,8,11 and 12). Although we observed a clear spatial transition of duration preferences in most of the participants (Fig. 5C, D), the orientations of this transition were much more variable compared to SMA (see S-Fig. 15B-C). Although this variability can reflect differences in the functional properties of these maps, it might be linked to the more complex morphology of these areas compared to SMA.

When we looked at the representation of 0.65 s in the two contexts (Fig. 6, see also S-Figs. 9 and 13 for individual maps), we observed very few vertices responding to 0.65 s independently from the context, and a spatial shift of vertices responding to 0.65 s according to the relative position of this duration within a temporal context. Differently from SMA, the spatial shift in these areas did not correlate with a shift in perception ( $r = 0.05$  for IPS;  $r = 0.07$  for auditory parabelt). Similar to what we observed in SMA, in IPS and parabelt areas, the cross-validation of the tuning preferences showed consistency of duration preferences across the fMRI runs (i.e., voxels keep their duration preferences across runs) and a gaussian-like tuning profile within each context, where the hemodynamic response peaks at the preferred duration and slowly decays with distance from it. Finally, in these brain regions we observed a clear segregation of duration preferences across the two contexts (see Fig. 7CD) i.e. voxels did not change their tuning across contexts. This last result suggests a key role of the temporal context in driving duration preferences and in shaping the maps. Also here, the maps in the two contexts are neither perfectly segregated nor overlapping (see Fig. 7A and S-Figs. 10 and 14).

At this point, to prove the robustness of the current findings we decided to analyze the fMRI data with a data driven approach. For this purpose, we ran a Multivariate Pattern Recognition Analysis (MVPA). The first goal was to check whether the six different durations in the two contexts could be predicted by the pattern of activity of the three ROIs of interest (i.e., SMA, IPS and parabelt area) but also by the activity of two task-unrelated control sites (i.e., Occipital pole -OP and Orbitofrontal cortex -OC). We decided to use

control ROIs to make sure that the pattern of activity observed in the ROIs of interest was specific to time processing in these areas. For MVPA, in each single subject we used the data of two fMRI runs (i.e., one for each context) to train a linear classifier on the 6 different durations and the remaining four fMRI runs (i.e., two for each context) to test the classification. This procedure was performed for all the possible combinations of training and testing runs (see Materials and Methods for details). For training and testing we used the beta values resulting from the GLM modelling of the offset of the different auditory durations (see Materials and Methods for more details). Figure 8A shows the results of the classification averaged across subjects, where for each ROIs is shown the confusion matrix with the classification accuracy of the three durations in the two contexts. The classification accuracy for each duration in the two contexts (i.e., the values in the diagonal of the matrix) was significantly above chance (0.17 is the chance level) in SMA and IPS compared to parabelt and the control ROIs (6 ROIs by 6 durations repeated measures ANOVA, ROI effect:  $F(4,5) = 6.25$   $p < 0.001$   $\eta^2 = .09$ ; SMA vs OP,  $t(5) = 7.28$   $p < 0.001$ ; IPS vs OP,  $t(5) = 5.34$   $p < 0.01$ ; parabelt vs OP,  $t(5) = 1.2$   $p = 0.14$ ), confirming, except for the parabelt areas, the previous model driven analysis. Moreover, the classification accuracy was significantly higher in SMA compared to IPS and parabelt areas (SMA vs IPS,  $t(5) = 3.37$   $p < 0.01$ ; SMA vs parabelt  $t(5) = 2.41$   $p < 0.05$ ; parabelt vs IPS  $t(5) = 0.68$   $p < 0.26$ ). This last result suggests that despite the similarity of patterns linked to duration preferences, duration selectivity seems more prominent in SMA compared to IPS and parabelt areas. This result was also replicated by performing the classification of the two contexts only (without considering the different durations, see S-Fig. 16 and Materials and Methods for more details). Next, to understand the contribution of the context in modulating duration preferences we ran a dissimilarity analysis, in which we correlated the betas associated with the offset of each duration (as resulting from the GLM analysis) within and across contexts. In SMA, IPS and parabelt the pattern of activity associated with the different durations was more similar within rather than across contexts (Fig. 8B). This result highlights the importance of the context in shaping the activity within these areas and in creating a relationship between durations belonging to the same context.

Finally, we ran an additional classification analysis in which we tried to predict the 6 different durations from the pattern of activity of the different duration selective clusters as previously identified with the GLM winner-take-all procedure. The MVPA was performed as before, using two fMRI runs for training (i.e., one for each context) and the remaining runs for testing. All possible combinations of training and testing runs were used. The result of this classification (see S-Fig. 17) shows that for each duration selective cluster, the majority of the voxels accurately predicts the duration originally preferred to that cluster. For example, in the 0.32 cluster, as defined by winner-take-all, there is a great number of voxels that are classified as preferring 0.32 s duration. This result thus confirms with a data driven approach the consistency of the duration preferences and proves the robustness of our original findings.

## Discussion

In this work we show the presence of a topographic representation of time in SMA, IPS and parabelt areas. Chronotopic maps were observed at the sound offset of different durations. Within these maps, duration selective voxels show a gaussian-like type of tuning, where response is enhanced for preferred

duration and slowly decays with distance from it. Maps were observed in different temporal contexts and although the cortical area covered by the different context maps was largely overlapping, distinct voxels responded to the different durations of the two contexts. Most importantly, different voxels responded to the same physical duration when this was perceived differently in the two contexts. Those clusters of voxels were also spatially shifted according to the relative position with each context and only in SMA this spatial shift correlated with the perceived difference of the same duration in the two contexts. MVPA analysis confirmed the presence in SMA and IPS of distinct patterns of activity for the different durations in the two contexts. However, these patterns were more easily detectable in SMA compared to IPS and parabelt. Finally, a dissimilarity analysis shows in all areas of interest a clear segregation of the activity associated with the different temporal contexts i.e., activity is more similar within rather than across contexts.

In humans, chronotopic maps and duration preferences have been described before in a wide network of brain regions including visual, parietal, premotor and prefrontal regions [6], [8], [7]. These maps have been described when participants were both passively viewing duration stimuli and when they were directly engaged in a duration discrimination task. However, only in SMA, these maps were linked to duration perception [6]. Here, differently from these previous studies, we show the presence of chronomaps for auditory stimuli and when the goal of duration sound encoding was a reproduction i.e., a motor task. Chronotopic maps were observed not only in high-level parietal and premotor brain regions as before, but also in sensory specific regions like the parabelt area. These data therefore suggest the presence of topographic representation of time across different stages of duration processing i.e., from auditory associative cortices to intraparietal sulcus to SMA. The redundancy of this temporal representation resembles the existing redundancy of spatial representations, where different brain areas host different spatial representations serving different functional purposes [13]. Our experimental design and the intrinsic spatial and temporal limitations of the fMRI technique do not allow us to specify the functional properties of these different maps. However, there are a few aspects of our results that might give a hint on the functional differences between these regions in duration encoding. The first is the significant difference between SMA, IPS and parabelt activity in predicting the different durations. Indeed, the MVPA analysis showed a progressive worsening of accuracy in predicting the different durations from the premotor to the sensory areas. Maps in IPS and in parabelt compared to SMA, also showed a high degree of inter-subject variability in orientation and, differently from SMA, the spatial shift observed for the duration shared between the contexts did not correlate with differences in perception. These three observations seem to suggest a special role of SMA in duration encoding. SMA, compared to IPS and parabelt areas, is decisively the area where durations must be read-out and recognized for forthcoming decisions. This interpretation is in line with previous results showing a correlation between SMA activity and duration perception [6], [14] and with the results of a recent effective connectivity study [8]. In this study the authors explored the connectivity architecture of five functionally distinct brain areas (i.e., cerebellum, primary visual cortex, IPS, SMA and Inferior frontal gyrus), significantly associated with the duration encoding of brief visual stimuli (ranging from 0.2 to 1 s). The results showed that the optimal effective connectivity model is the one in which the cerebellum has feedback and/or feedforward

connections from and to all other network nodes. SMA is the only area that, while being modulated by the activity of cerebellum, IPS and V1, does not influence the activity of any other brain region [8]. According to this work SMA seems to be the ultimate stage of duration recognition, whereas IPS is the area whose activity is greatly affected by the incoming duration information (is the area sensitive to the duration input). In light with our current and previous works we can therefore hypothesize that duration information is first extracted in auditory regions and then passed to IPS where a first reading of temporal signals occurs (i.e., “duration input” area) and from IPS duration information reaches SMA, the final stage of duration recognition, where duration will be ready for decision. In humans, the role of both SMA and inferior parietal lobule in temporal perception has been extensively documented [15],[16],[17]. Both areas have been implicated in a variety of timing tasks [15], [18],[19] with a range of durations spanning from a few hundreds of milliseconds to a few seconds [20],[21] and with stimuli of different sensory modalities [22],[23]. It is therefore likely that both areas constitute the core of the timing network.

Compared to previous studies here we were able to specify a few important properties of the maps.

The observation that the same physical duration engaged the activation of different voxels when perceived differently in the two contexts, and that these voxels are spatially shifted according to the relative position of this duration within each context, seems to suggest that time is mapped in a relative fashion. However, the maps in the two contexts are not perfectly relative, since first, distinct voxels are active for the different durations in the two contexts. Second, there is no remapping of the tuning across contexts i.e., within a given duration selective cluster of voxels, the BOLD signal does not peak for durations sharing the same position within a distribution. Third, maps are strongly modulated by the context, as shown by the dissimilarity analysis within each ROIs. The activity pattern is more similar within the durations of a context rather than across them. Even though the clusters of voxels active in the two contexts are not spatially segregated, they largely overlap on the cortical surface.

The observation that voxels change their duration preferences according to the context and the position of the durations within it, seems to suggest a certain degree of flexibility of duration representation in these maps. Even though any comparison with tuning mechanisms explored at electrophysiological level in single-cells has to be taken with caution, we believe that our data are compatible with some of the basic properties of other existing topographic maps. The flexibility observed in chrononompas for example, has also been reported in “sensory” maps, where the tuning to a specific stimulus feature (e.g., orientation, spatial and temporal frequency, motion direction in the visual domain) change after perceptual adaptation [24],[25],[26]. A change of response preferences measured with fMRI, has also been observed in numerosity maps after perceptual adaptation [27]. And similarly to “sensory” maps, where this flexibility is limited, adaptation effects for example, occur only when there is an optimal distance between adaptor stimulus and test, here there is no total reshape of the two contexts maps but different voxels respond to the different durations in the contexts. There is an important caveat to make when comparing chrononompas to other more “sensory” maps. Time maps together with numerosity maps have never been described in primary sensory cortices [7], [6], probably because time, like numerosity, lacks a proper “sensory receptive space”. It is therefore plausible that these maps reflect a high-level stage of

temporal processing. Low level sensory areas, like primary visual cortex, for example, are indeed sensitive to changes of stimulus duration [28], but this sensitivity is reflected in the change of the hemodynamic response amplitude i.e., a sublinear increase of BOLD with increasing stimulus duration, and not in a tuning-like response [29]. This difference seems to suggest that duration preferences arise later in the temporal information processing stream, perhaps as a result of the integration of the sensory drive that comes from primary sensory areas.

The modulation of stimulus contexts in shaping the maps is also an interesting and novel aspect of our findings. Although the effect of temporal context has been well documented at behavioral level [30], [31], only very few studies [11], [32] have explored the neural signature of this effect.

A recent EEG study, for example, using a very similar auditory temporal reproduction task, has shown that temporal context affects the neural dynamics during the encoding of the stimulus duration. Specifically, longer previous durations decrease CNV and P2 amplitude and increase beta power [11], suggesting, similarly to our results, a modulation of temporal context on perceptual rather than memory processing.

In summary in this work we show the existence of chronomaps across auditory, parietal, and premotor regions. In SMA and IPS chronomaps are sensory modality and task independent. All maps show a high degree of flexibility with different voxels responding to the same physical duration (i.e., 0.65 s) in the two temporal contexts; these voxels are spatially shifted according to the relative position of this duration within the context. This flexibility though is not absolute, voxels do not change their duration preferences across contexts, but more often different voxels are active for the different durations in the two contexts. The temporal context seems indeed very powerful in making the pattern of activity associated with the different durations more similar within a context rather than across contexts.

Overall these results suggest that time is represented in the maps in a partially relative fashion and that the temporal contexts play a pivotal role in determining duration preferences.

## Declarations

## Acknowledgments

Financial support has been provided by the European Research Council -ERC (Grant Agreement No 682117, BiT-ERC-2015-CoG) and from the Italian Ministry of University and Research under the call FARE (project ID: R16X32NALR) to D.B. We thank Mr. Gianfranco Fortunato for his feedback on an earlier version of the manuscript.

## References

1. Paton JJ, Buonomano D V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron*. 2018;98: 687–705. doi:10.1016/j.neuron.2018.03.045

2. Merchant H, Harrington DL, Meck WH. Neural Basis of the Perception and Estimation of Time. *Annu Rev Neurosci.* 2013/06/04 . 2013;36: 313–336. doi:10.1146/annurev-neuro-062012-170349
3. Mita A, Mushiake H, Shima K, Matsuzaka Y, Tanji J. Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat Neurosci.* 2009;12: 502–507. doi:10.1038/nn.2272
4. Merchant H, Perez O, Zarco W, Gamez J. Interval tuning in the primate medial premotor cortex as a general timing mechanism. *J Neurosci.* 2013/05/24. 2013;33: 9082–9096. doi:10.1523/JNEUROSCI.5513-12.201333/21/9082 [pii]
5. Gouvêa TS, Monteiro T, Motiwala A, Soares S, Machens C, Paton JJ. Striatal dynamics explain duration judgments. *Elife.* 2015;4. doi:10.7554/eLife.11386
6. Protopapa F, Hayashi MJ, Kulashekhar S, Zwaag W van der, Battistella G, Murray MM, et al. Chronotopic maps in human supplementary motor area. *PLOS Biol. Public Library of Science;* 2019;17: e3000026. doi:10.1371/JOURNAL.PBIO.3000026
7. Harvey BM, Dumoulin SO, Fracasso A, Paul JM. A Network of Topographic Maps in Human Association Cortex Hierarchically Transforms Visual Timing-Selective Responses. *Curr Biol.* 2020; doi:10.1016/j.cub.2020.01.090
8. Protopapa F, Hayashi MJ, Kanai R, Buetti D. Topographic Connectivity in a Duration Selective Cortico-Cerebellar Network. *bioRxiv.* 2020; 2020.04.08.031385. doi:10.1101/2020.04.08.031385
9. Jazayeri M, Shadlen MN. Temporal context calibrates interval timing. *Nat Neurosci.* 2010;13: 1020–1026. doi:10.1038/nn.2590
10. Maaß SC, Schlichting N, van Rijn H. Eliciting contextual temporal calibration: The effect of bottom-up and top-down information in reproduction tasks. *Acta Psychol (Amst). Netherlands;* 2019;199: 102898. doi:10.1016/j.actpsy.2019.102898
11. Damsma A, Schlichting N, van Rijn H. Temporal Context Actively Shapes EEG Signatures of Time Perception. *J Neurosci.* 2021;41: 4514 LP – 4523. doi:10.1523/JNEUROSCI.0628-20.2021
12. Glasauer S, Shi Z. The origin of Vierordt’s law: The experimental protocol matters. *PsyCh J. John Wiley & Sons, Ltd;* 2021;n/a. doi:https://doi.org/10.1002/pchj.464
13. Derdikman D, Moser EI. A manifold of spatial maps in the brain. *Trends in Cognitive Sciences.* 2010. doi:10.1016/j.tics.2010.09.004
14. Coull JT, Charras P, Donadieu M, Droit-Volet S, Vidal F. SMA Selectively Codes the Active Accumulation of Temporal, Not Spatial, Magnitude. *J Cogn Neurosci. MIT PressOne Rogers Street, Cambridge, MA 02142-1209USAjournals-info@mit.edu;* 2015;27: 2281–2298. doi:10.1162/jocn\_a\_00854
15. Wiener M, Turkeltaub P, Coslett HB. The image of time: a voxel-wise meta-analysis. *Neuroimage.* 2010;49: 1728–1740.

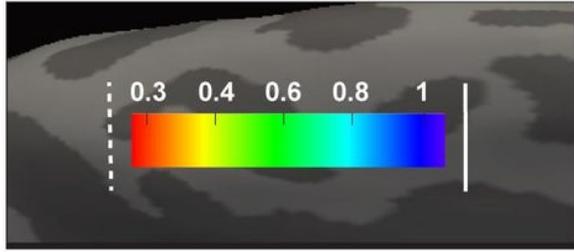
16. Hayashi MJ, Ditye T, Harada T, Hashiguchi M, Sadato N, Carlson S, et al. Time Adaptation Shows Duration Selectivity in the Human Parietal Cortex. Zatorre R, editor. *PLOS Biol. Public Library of Science*; 2015;13: e1002262. doi:10.1371/journal.pbio.1002262
17. Hayashi MJ, van der Zwaag W, Buetti D, Kanai R. Representations of time in human frontoparietal cortex. *Commun Biol.* 2018;1: 233. doi:10.1038/s42003-018-0243-z
18. Buetti D, Macaluso E. Auditory temporal expectations modulate activity in visual cortex. *Neuroimage.* 2010;51. doi:10.1016/j.neuroimage.2010.03.023
19. Buetti D, Walsh V, Frith C, Rees G. Different brain circuits underlie motor and perceptual representations of temporal intervals. *J Cogn Neurosci.* 2008;20. doi:10.1162/jocn.2008.20017
20. Morillon B, Kell CA, Giraud AL. Three stages and four neural systems in time estimation. *J Neurosci.* 2009;29: 14803–14811.
21. Lewis PA, Miall RC. Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia.* 2003;41: 1583–1592.
22. Pastor MA, Day BL, Macaluso E, Friston KJ, Frackowiak RS. The functional neuroanatomy of temporal discrimination. *J Neurosci.* 2004;24: 2585–2591.
23. Coull JT, Nazarian B, Vidal F. Timing, storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. *J Cogn Neurosci.* 2008;20: 2185–2197.
24. Kohn A, Movshon JA. Adaptation changes the direction tuning of macaque MT neurons. *Nat Neurosci.* 2004;7: 764–772. doi:10.1038/nn1267
25. Kohn A. Visual adaptation: Physiology, mechanisms, and functional benefits. *J Neurophysiol.* 2007;97: 3155–3164. doi:10.1152/jn.00086.2007
26. Dragoi V, Sharma J, Sur M. Adaptation-Induced Plasticity of Orientation Tuning in Adult Visual Cortex. *Neuron.* 2000;28: 287–298. doi:https://doi.org/10.1016/S0896-6273(00)00103-3
27. Tsouli A, Cai Y, van Ackooij M, Hofstetter S, Harvey BM, te Pas SF, et al. Adaptation to visual numerosity changes neural numerosity selectivity. *Neuroimage.* 2021;229: 117794. doi:https://doi.org/10.1016/j.neuroimage.2021.117794
28. Zhou J, Benson NC, Kay KN, Winawer J. Compressive temporal summation in human visual cortex. *J Neurosci.* 2018; doi:10.1523/JNEUROSCI.1724-17.2017
29. Hendrikx E, Paul J, Ackooij M van, Stoep N van der, Harvey B. Derivation of visual timing-tuned neural responses from early visual stimulus representations. *Nat Portf.* 2021; doi:10.21203/rs.3.rs-579436/v1
30. Roach NW, McGraw P V, Whitaker DJ, Heron J. Generalization of prior information for rapid Bayesian time estimation. *Proc Natl Acad Sci U S A. National Academy of Sciences*; 2017;114: 412–417. doi:10.1073/pnas.1610706114
31. Lejeune H, Wearden JH. Vierordt's the Experimental Study of the Time Sense (1868) and its legacy. *Eur J Cogn Psychol.* 2009;21: 941–960. doi:10.1080/09541440802453006

32. Murai Y, Yotsumoto Y. Context-dependent neural modulations in the perception of duration. *Front Integr Neurosci*. 2016;10: 1–11. doi:10.3389/fnint.2016.00012
33. Brainard DH. The Psychophysics Toolbox. *Spat Vis*. 1997;10: 433–436. doi:10.1163/156856897X00357
34. van der Zwaag W, Francis S, Head K, Peters A, Gowland P, Morris P, et al. fMRI at 1.5, 3 and 7 T: characterising BOLD signal changes. *Neuroimage*. 2009/05/19. 2009;47: 1425–1434. doi:S1053-8119(09)00500-X [pii]10.1016/j.neuroimage.2009.05.015
35. Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J. Classical and Bayesian inference in neuroimaging: applications. *Neuroimage*. 2002;16: 484–512.
36. Fischl B, Salat D ~ H., Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002;33: 341–355.
37. Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*. 2010;53: 1–15. doi:10.1016/j.neuroimage.2010.06.010
38. Oosterhof NN, Connolly AC, Haxby J V. CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Front Neuroinform*. 2016;10: 27. doi:10.3389/fninf.2016.00027
39. Chang C-C, Lin C-J. LIBSVM: A Library for Support Vector Machines. *ACM Trans Intell Syst Technol*. New York, NY, USA: Association for Computing Machinery; 2011;2. doi:10.1145/1961189.1961199
40. Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. *Proc Natl Acad Sci U S A*. 2006;103: 3863–3868. doi:10.1073/pnas.0600244103
41. Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci*. 2008;2: 4. doi:10.3389/neuro.06.004.2008

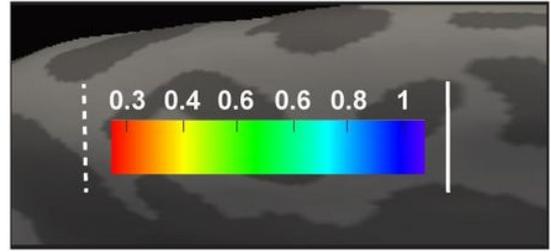
## Figures

A

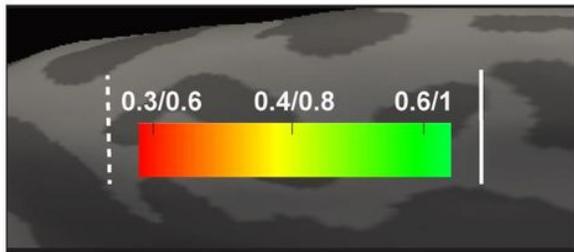
1. Absolute map



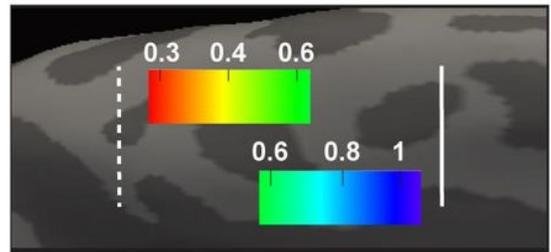
2. Absolute map modulated by context



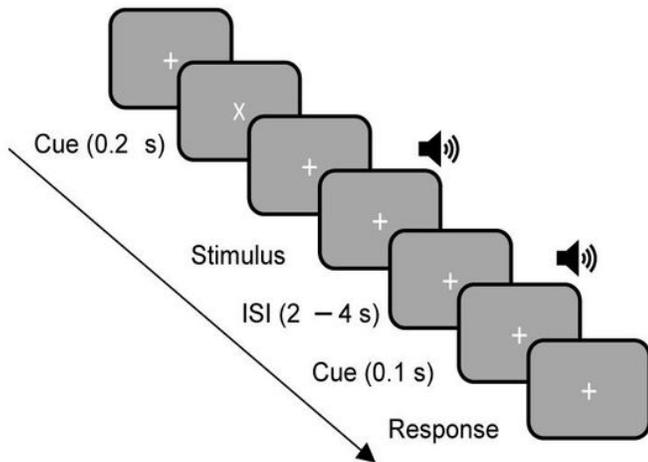
3. Relative map



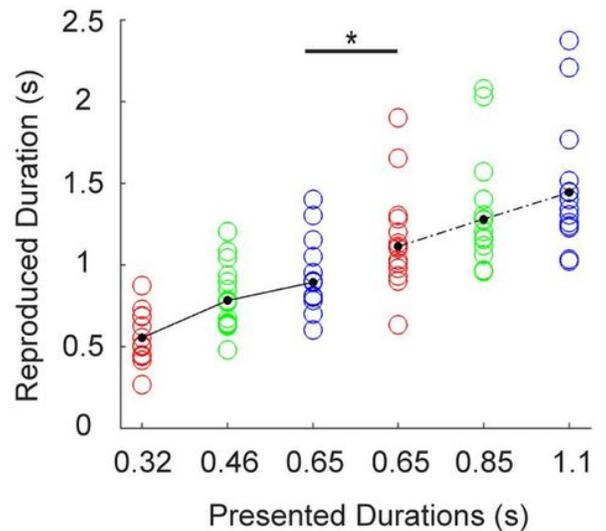
4. Relative map modulated by context



B



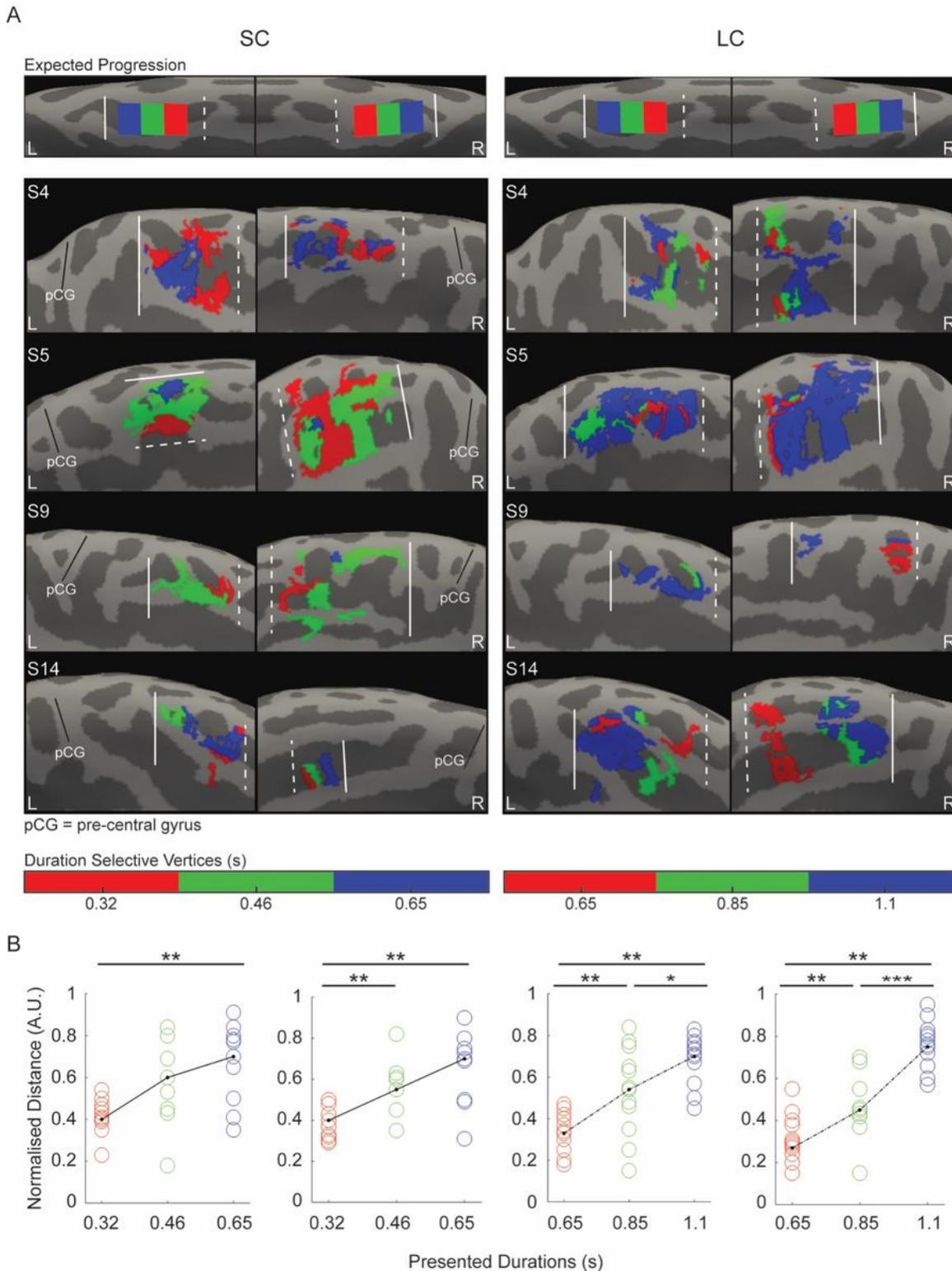
C



**Figure 1**

Predictions and experimental paradigm with behavioral results. (A) Sketch of four possible duration representations in chromamaps associated with changes in perception and temporal contexts. Lines are hypothetical maps' borders. The dashed line is the shortest edge of the map. The different panels depict the four predictions described in the text. (B) Schematic representation of the stimuli in a trial. Each trial began with a visual cue (a fixation cross changing briefly its shape), after 1.3–2.3 s a pure tone (10KHz in pitch) of a given duration was played through headphones. In the short context the sound could be played for 0.32, 0.46, or 0.65 s, in the long context it could be played for 0.65, 0.85 or 1.1 s. After a variable interval ranging from 2 to 4 seconds, a burst of white noise lasting 0.1s, instructed the subject to

reproduce the previous stimulus duration with a key press. (C) Group average (N = 14) of the reproduced stimulus durations plotted separately for each of the contexts. The colored circles are the different subjects; the black dots are the mean of the distribution. \*p = 0.04.

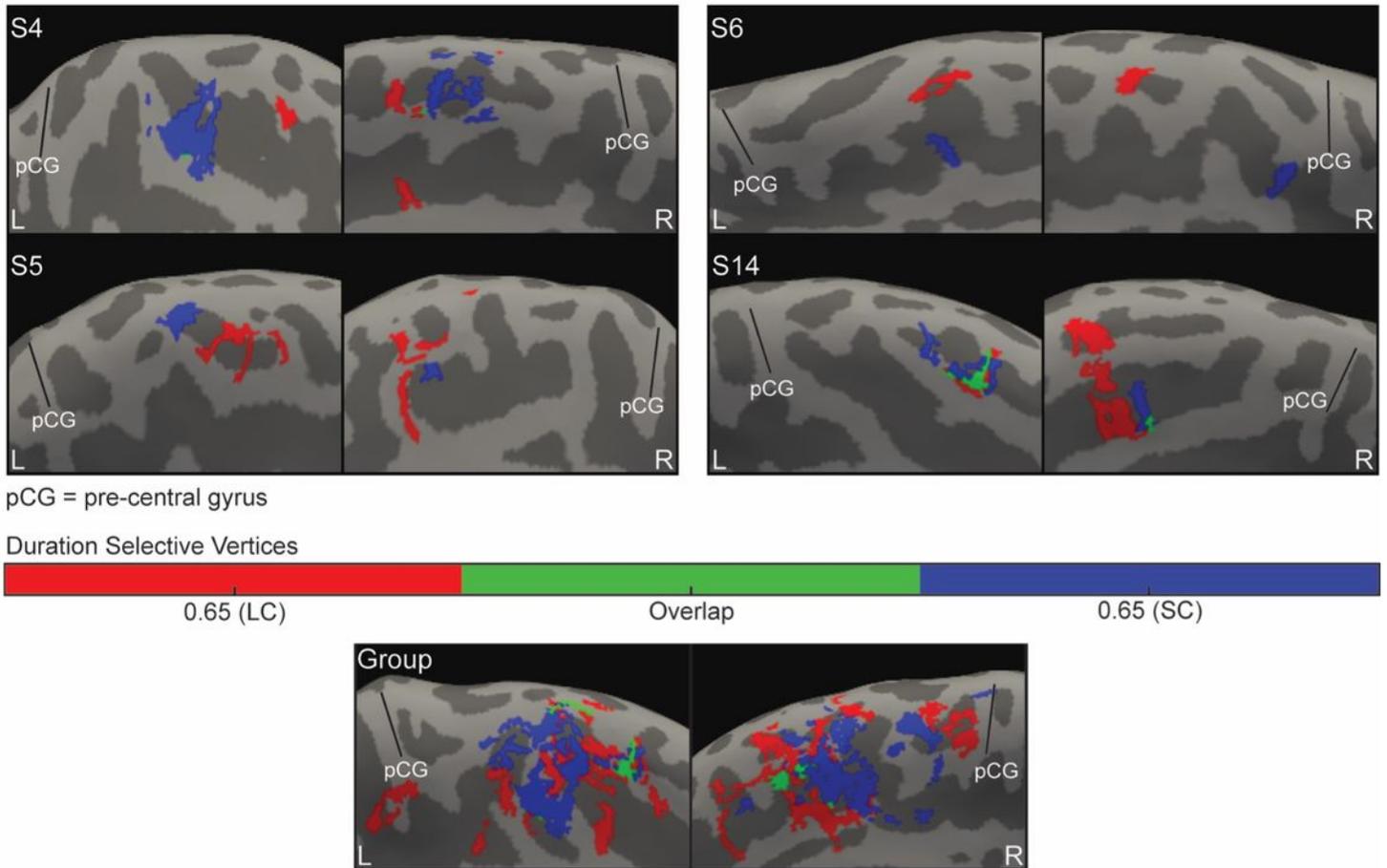


**Figure 2**

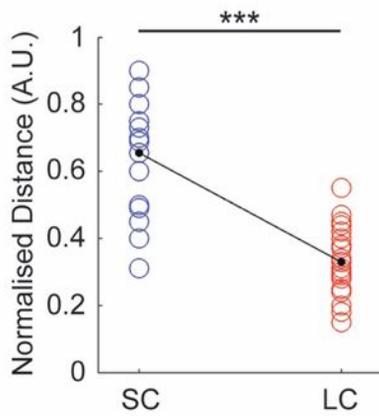
Auditory chronomaps in SMA. (A) Medial view of left (L) and right (R) hemispheres of four individual subjects. Overlaid on the individual inflated cortical surface are clusters of vertices classified with a

winner-take-all procedure as maximally responsive to the three durations in the short (SC, 0.32, 0.46, and 0.65 s) and long (LC, 0.65, 0.85, and 1.1 s) context. In red, green and blue are the clusters preferring respectively the shortest, the intermediate and the longest duration of the context. The white dashed line is the border at the shortest edge of the map (closest to the shortest duration), the solid line the border at the longest edge of the map (closest to the longest duration of the range). pCG is precentral gyrus. (B) The plots show for each hemisphere and context the distance of the different duration selective clusters (i.e., the average of the vertices in a cluster), color-coded as in (A), from the shortest border of the map (dashed line). The distances were computed in each individual map and were normalized to the individual map size (distance between short and long edges of the maps, see Material and Methods for details). The colored circles are the individual data, the lines represent the group average (continuous line is SC, dashed line is LC). The normalized distances were computed for each context and each individual subject on chronomaps overlaid on flattened surfaces in the participant's native space. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

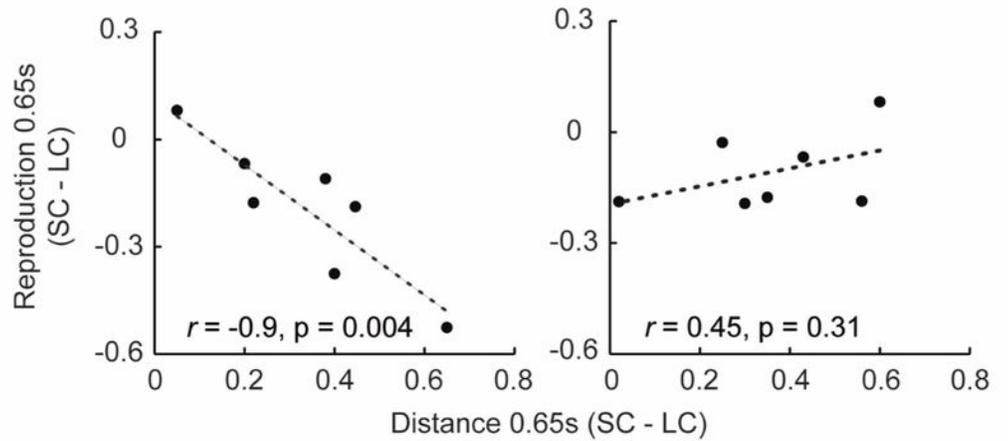
A



B



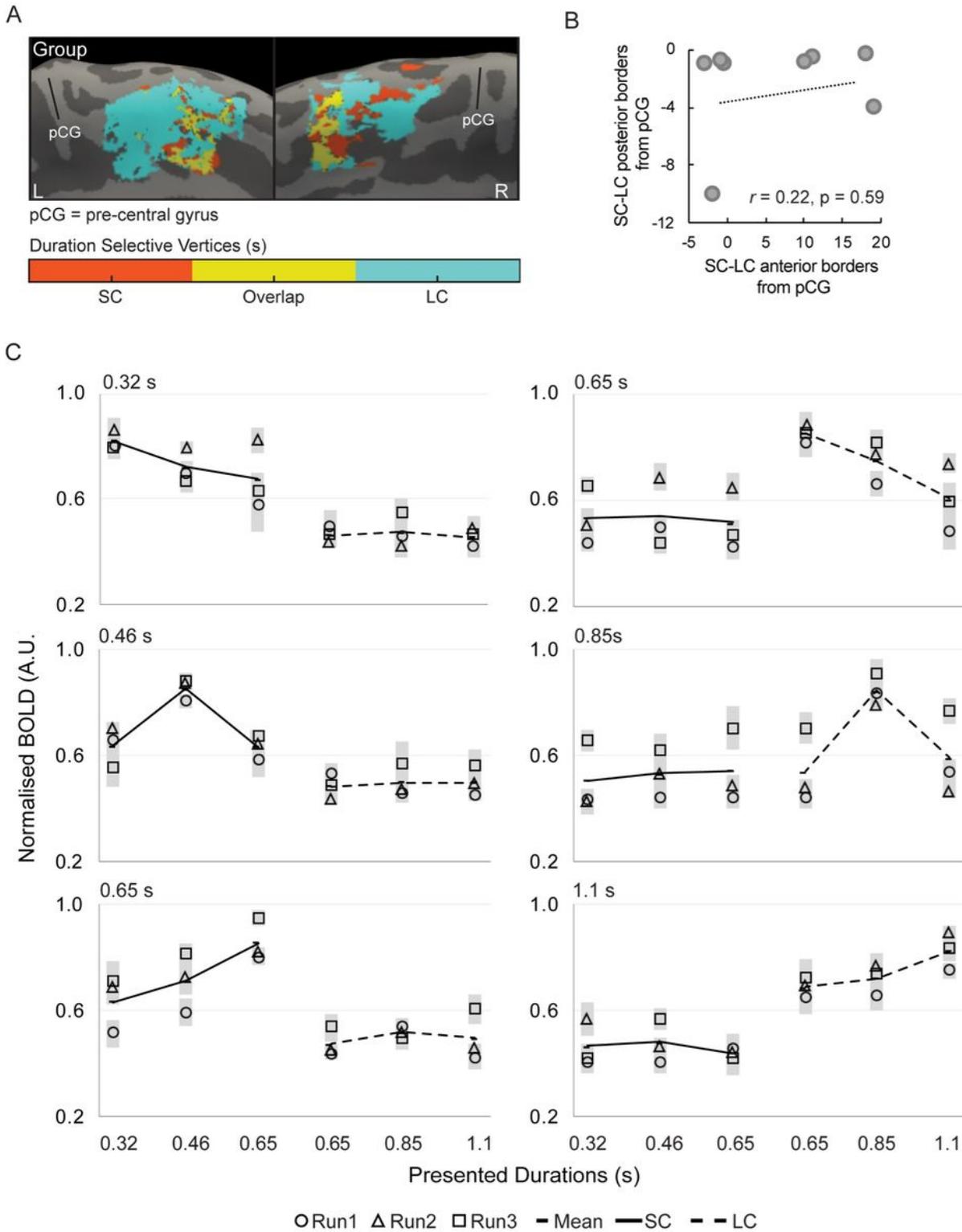
C



**Figure 3**

Representation of the duration shared between contexts in SMA. (A) Medial view of left (L) and right (R) hemispheres of four individual subjects and of the group. Overlaid on the subject's inflated cortical surface are clusters of vertices classified with a winner-take-all procedure as maximally responsive to 0.65 s when presented in short context (SC, in blue) and long context (LC, in red). In green are the clusters responsive to 0.65 s in both the contexts. Individual maps are overlaid on individual brains in native

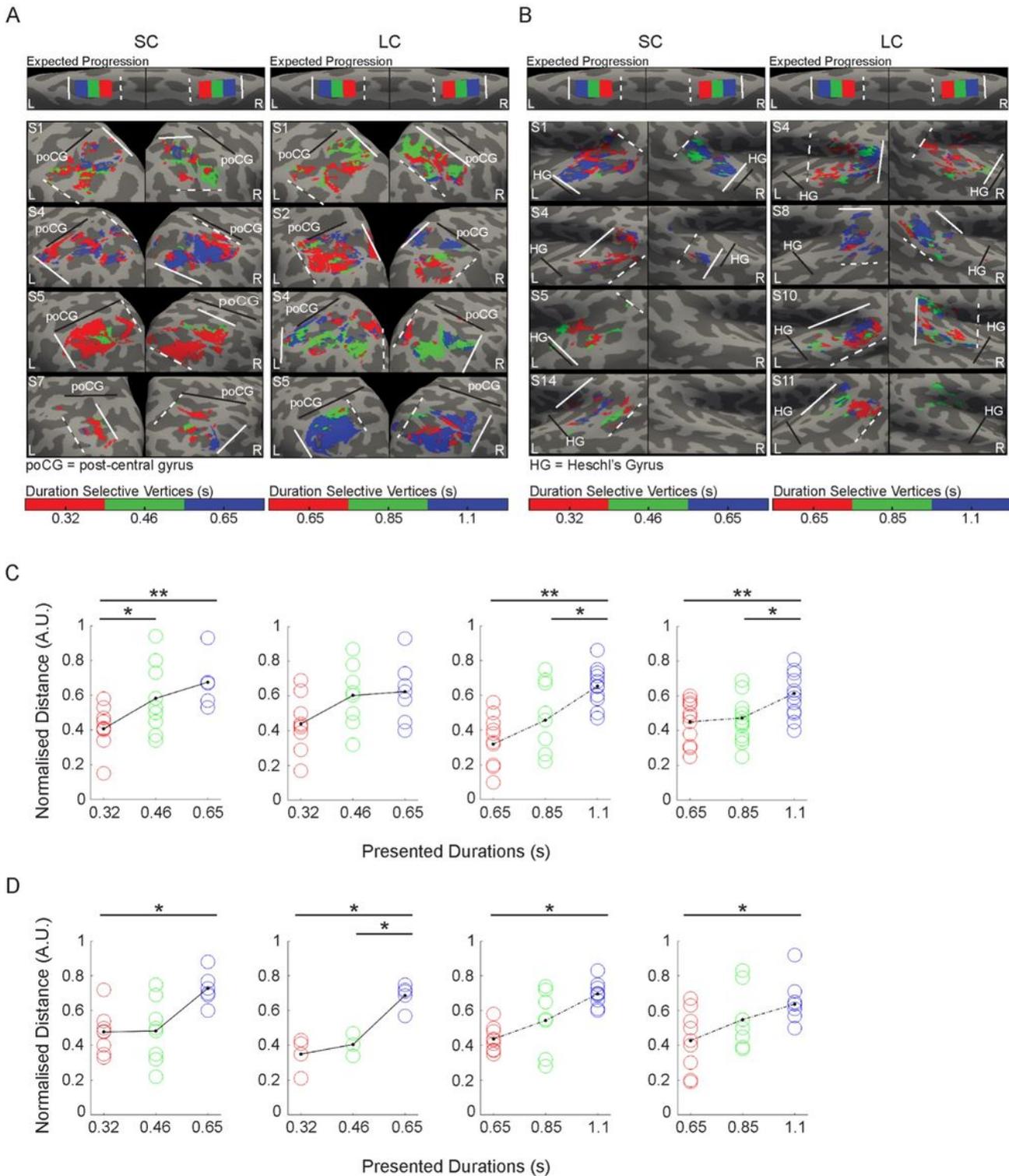
space. For the group image (at the bottom of the panel) the clusters of the 14 subjects were morphed onto a common inflated surface to form a group map. (B) Shows the normalized distance of the 0.65 s duration selective clusters in SC and LC. In red the clusters of the LC, in blue those in the SC, for simplicity, the data of the right and left hemispheres are considered together ( $N(\text{SC}) = 15$  and  $N(\text{LC}) = 21$ ). The circles are individual subjects/hemispheres, the black points their means. Each circle is the average distance of all vertices in a 0.65 s cluster. The normalized distances were computed in each subject, hemisphere and context separately as the difference (normalized to the map's size) from the shortest edge of the map.  $***p < 0.001$  (C) Shows the correlation between the difference in reproduction time (time from key-press to key release) of 0.65 s (y axis) and the normalized distance of the 0.65 s clusters in short and long contexts. Data points are individual subjects and the dashed line is the best fit line. In the leftward panel is the left hemisphere, in the rightward is the right hemisphere. For the correlation, we considered for each hemisphere only the subjects that had significant activation of the 0.65 s cluster in the two contexts ( $N=7$ ).



**Figure 4**

Overlap between the maps in the two contexts and duration tuning in SMA. (A) Shows the overlap between the maps of the two contexts. Medial view of left (L) and right (R) hemispheres of the group, with the clusters corresponding to the short context in orange, those corresponding to the long context in pale blue and their overlap in yellow. Here we show morphed onto a common inflated cortical surface only the individual chromomaps where the progression was in the anterior to posterior orientation in the

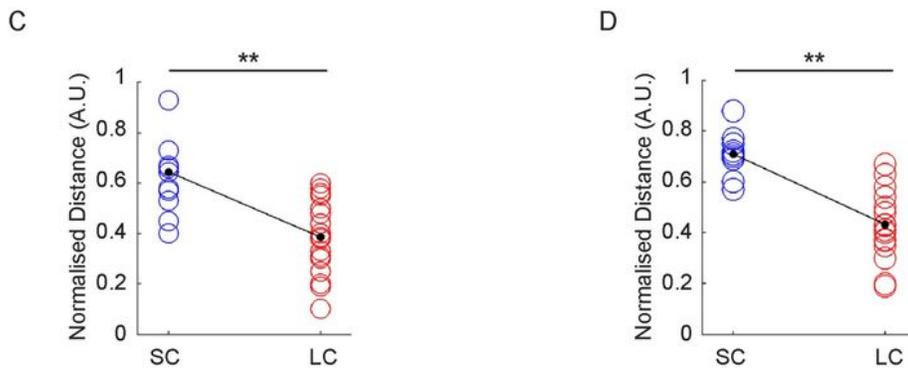
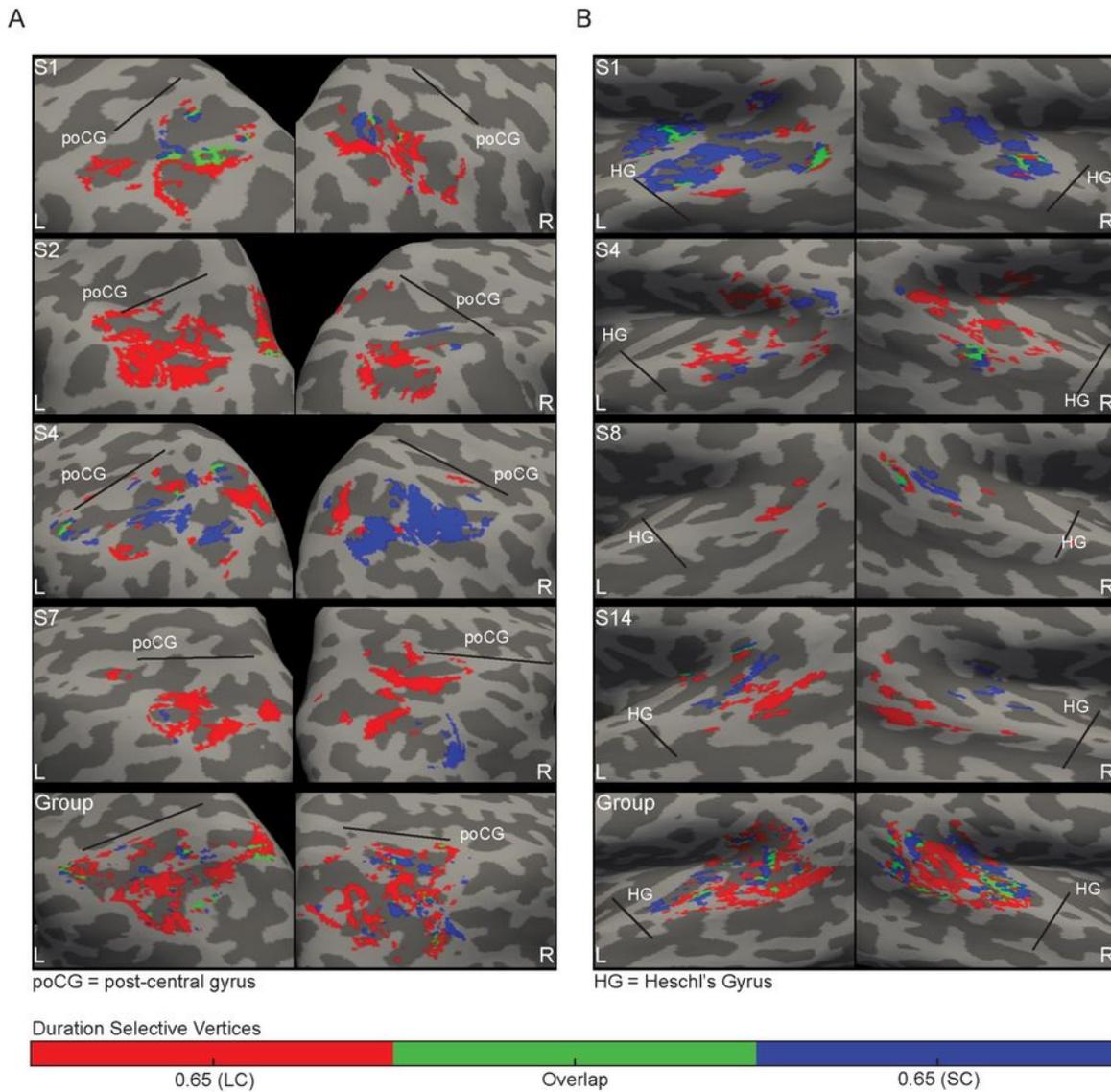
two contexts in each hemisphere ( $N(\text{left}) = 5$ ,  $N(\text{right}) = 3$ ) see Material and Methods for more details). pCG is precentral gyrus. (B) differences between short (SC) and long (LC) context of anterior (x axis) and posterior (y axis) borders from pCG measured in each individual subject (each dot is a subject). Only maps in the anterior-to-posterior orientation were included ( $N=8$ ). Before subtracting the LC from the SC context border, for each individual border the distance from the precentral gyrus was also computed (see Materials and Methods for more details). (C) Shows the duration tuning of the different duration selective clusters of voxels within and across contexts computed using a cross-validation approach across fMRI sessions (leave one-run-out approach). Group average of the normalized BOLD responses (y-axis) of the different duration-selective voxels (in different panels) for preferred and non-preferred durations within and across contexts in the different fMRI sessions (diamonds, squares and circles). On the x-axis are the six different durations in the two contexts. The lines are the average across fMRI sessions (continuous line for SC, dashed line for LC). The BOLD signal in the duration-selective voxels is aligned to the presentation timings of the different durations (i.e., second volume after duration offset, see Materials and Methods for more details). The duration selective clusters were identified from a single run of the appropriate context (one for each context) and the normalized BOLD signal was extracted from the same clusters in the remaining runs. Normalization was performed in each individual subject to the mean signal intensity of the appropriate fMRI run (see Materials and Methods for more details).  $m$  is the value of the estimated slope.



**Figure 5**

Auditory chromomaps in IPS and parabelt areas. (A, B) Medial view of left (L) and right (R) hemispheres of four individual subjects. Overlaid on the individual inflated cortical surface are clusters of vertices classified with a winner-take-all procedure as maximally responsive to the three durations in the short (SC, 0.32, 0.46, and 0.65 s) and long (LC, 0.65, 0.85, and 1.1 s) context in (A) maps in IPS and in (B) maps in parabelt areas. In red, green and blue are the clusters preferring respectively the shortest, the intermediate

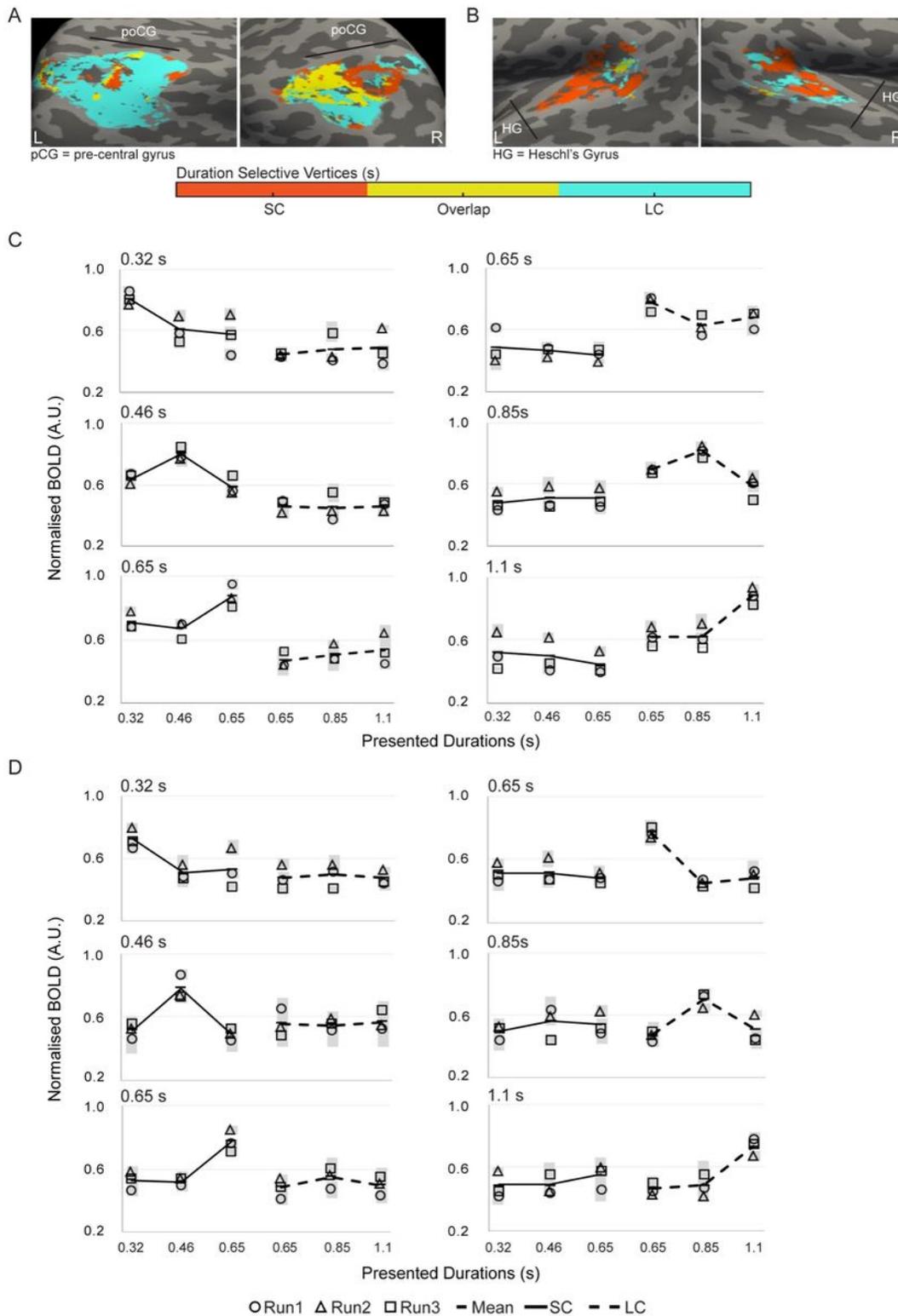
and the longest duration of the context. The white dashed line is the border at the shortest edge of the map (closest to the shortest duration of the range), the solid line the border at the longest edge of the map (closest to the longest duration). poCG is postcentral gyrus; HG is Heschl's Gyrus. (C, D) the plots show for each hemisphere and context the distance of the different duration selective clusters color coded as in (A, B) from the shortest border of the map (dashed line). The distances were computed in each individual map and were normalized to the individual map size (distance between short and long edges of the maps, see Material and Methods for details). The colored circles are the individual data, the lines and the black dots represent the group average (continuous line is SC, dashed line is LC). Each disk in the plot is the average distance of the vertices in a given cluster. The normalized distances were computed for each context and each individual subject on chronomaps overlaid on flattened surfaces in the participant's native space. \* $p < 0.05$ , \*\* $p < 0.01$ .



**Figure 6**

Representation of the duration shared between contexts in IPS and parabelt areas (A, B) Medial view of left (L) and right (R) hemispheres of four individual subjects and of the group. Overlaid on the subject's inflated cortical surface are clusters of vertices classified with a winner-take-all procedure as maximally responsive to 0.65 s when presented in SC (blue) and LC (red). In green are the clusters responsive to 0.65 s in both the contexts. In (A) the IPS, in (B) parabelt areas. Individual maps are overlaid on individual

brains in native space. For the group image (at the bottom of the panel) the clusters of the 14 subjects were morphed onto a common inflated surface to form a group map. (C, D) Show the normalised distance of the 0.65 s duration selective vertices in SC and LC. For simplicity, the data of the right and the left hemisphere are considered together (IPS, N = 10 (SC), 20 (LC); parabel, N = 9 (SC), 16 (LC)). The circles are individual subjects/hemispheres, the black dots their means. The normalized distances were computed in each subject, hemisphere and context separately as the difference from the shortest edge normalized to the map size. Each disk in the plot is the average distance of the vertices in a 0.65 s cluster, in (C) the IPS, in (D) parabelt areas. \*\*p < 0.01

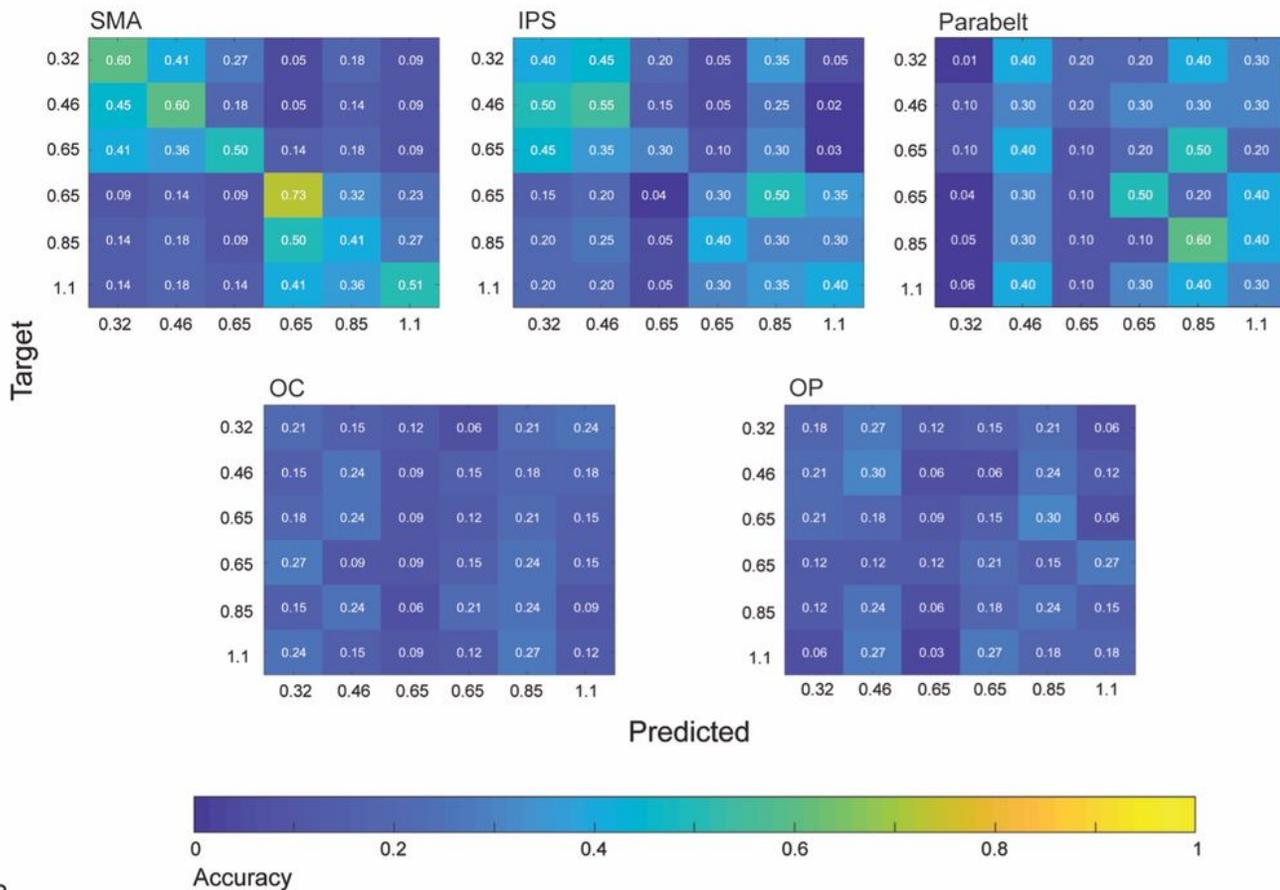


**Figure 7**

Overlap between the maps in the two contexts and duration tuning in IPS and parabelt areas. In (A) and (B) maps of the two contexts in IPS and parabelt areas. (A, B) Medial view of left (L) and right (R) hemispheres of the group, with the clusters corresponding to the short context in orange, those corresponding to the long context in pale blue and their overlap in yellow. Here we show morphed onto a common cortical inflated brain only the individual chromomaps where the progression was in the same

orientation in the two contexts and in each hemisphere (90° orientation in IPS N= 7 subjects and 180° orientation in parabelt areas N=3 subjects), see Material and Methods for more details). SC, short context; LC, long context; pCG, precentral gyrus. (C, D) Show the duration tuning of the different duration selective clusters within and across contexts computed using a cross-validation approach across fMRI sessions (leave one-run-out approach). In (C) IPS in (D) parabelt areas group average of the normalized BOLD responses (y-axis) of the different duration-selective voxels (in different panels) for preferred and non-preferred durations within and across contexts in the different fMRI sessions (diamonds, squares and circles). On the x-axis are the six different durations in the two contexts. The lines are the average across fMRI sessions (continuous line for SC, dashed line for LC). The BOLD signal in the duration-selective voxels is aligned to the presentation timings of the different durations (i.e., second volume after duration offset, see Materials and Methods for more details). The duration selective clusters were identified from two runs of the appropriate context (one for each context) and the normalized BOLD signal was extracted from the same clusters in the remaining runs. Normalization was performed in each individual subject to the mean signal intensity of the appropriate fMRI run (see Materials and Methods for more details).  $m$  is the value of the estimated slope.

A



B

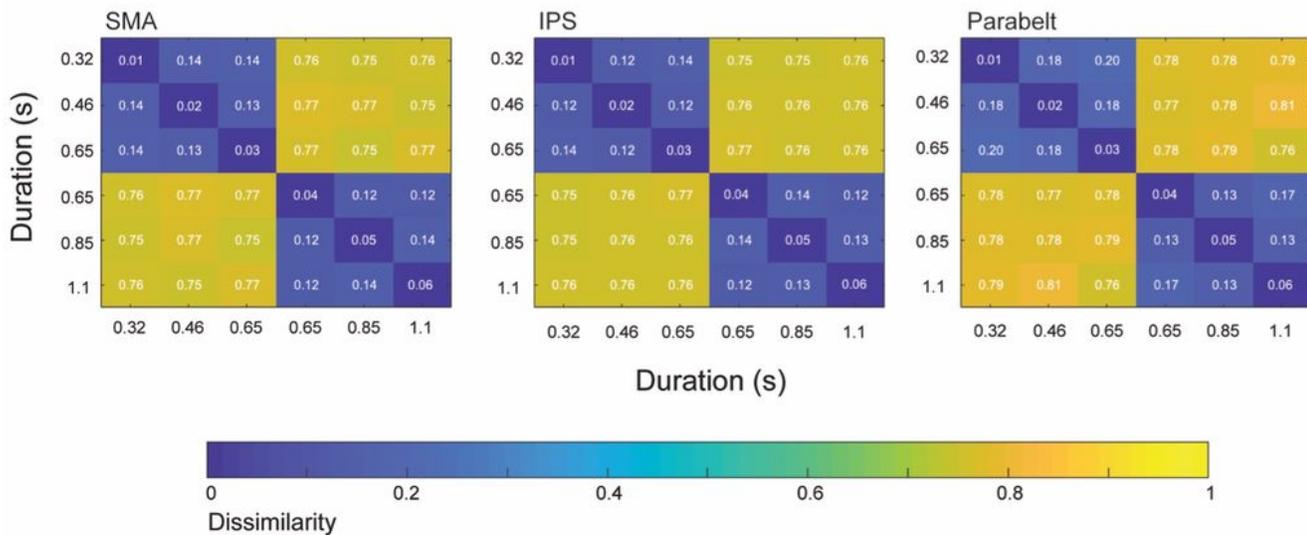


Figure 8

Decoding time with MVPA and representation similarity analysis. (A) Results of the MVPA analysis using a leave-one-run-out approach. (A) shows the mean decoding accuracy of the different durations in the two contexts from SMA, IPS, auditory parabelt, orbitofrontal cortex (OC) and occipital pole (OP) activity. For each subject and in each ROI, a support vector machine (SVM) classifier was trained to classify the pattern of activity associated with the six durations in two runs (x-axis). The classifier was then tested

using response patterns from the left-out run (y-axis). The overall classification accuracy was then computed by averaging the classification accuracy from all iterations. A classification accuracy at chance level was equal to 0.17 (1/6). In the confusion matrices here is reported the average decoding accuracy across subjects. (B) The results from a multivariate representational dissimilarity analysis, showing the mean dissimilarity between the neural activity associated with each duration in the two contexts. For each subject and for each ROI, a fMRI Representational Dissimilarity Matrix (RDM) was computed through a pair-wise (dis)similarity ( $1 - \text{Pearson's } r$ ) between each combination of the 6  $\beta$  weights associated with the offset of the 6 durations. The warmer the color, the greater the dissimilarity. In the matrices are reported the average dissimilarity coefficient across subjects.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterialKulashekharetal.v04.docx](#)