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

Shrikanth Kulashekhar
International School for Advanced Studies (SISSA)
Gianfranco Fortunato
International School for Advanced Studies (SISSA)
Sarah Maass
University of Groningen
Hedderik van Rijn
University of Groningen
Domenica Bueti (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: September 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-753370/v3

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 (Merchant, Harrington, et al., 2013; Paton & Buonomano, 2018).

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 (Gouvêa et al., 2015; Merchant, Pérez, et al., 2013; Mita et al., 2009). In humans these units are also topographically organized on the cortical surface to form chronomaps (Protopapa et al., 2019; Harvey et al., 2020). Chronotopic maps associated with visual temporal discrimination tasks have been observed in the Supplementary Motor Area (SMA Protopapa et al., 2019, 2020), 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 (Harvey et al., 2020).

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

(see for instance van Ackooij et al., 2022) 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 1.1s, 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 (Jazayeri & Shadlen, 2010; Maaß et al., 2019). 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 (Glasauer & Shi, 2021; Jazayeri & Shadlen, 2010). Although this effect of regression to the mean has been extensively documented at a behavioral level, and recent EEG studies have focussed on traces of the Bayesian integration (Damsma et al., 2021), 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 categorial, *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.

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 1.1s. The presentation of the sound 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 $F(1,2) = 63.3$ $p < 0.001$ 0 $\eta 2 = 1.01$). Overall the mean reproduction in the shortest context was shorter (0.76 s ± 0.24 s; mean ± standard deviation) compared to the longest context (1.4 s ± 0.4 s; mean ± 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 $t(13) = -2.16$, $p = 0.04$); 0.89 s ± 0.23 s in the short context and 1.11 s ± 0.31 s in the long context, confirming that the same physical duration (0.65s) 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). 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. The result of this contrast revealed differences in the activation of the auditory parabelt areas, the intraparietal sulcus (IPS), and the Supplementary Motor Area (SMA, see S-Fig.1).

We then focused on each of these regions 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, (i.e., the contrast of interest for each sound and context was sound offset - response onset, the contrast was $p_{FWE} < 0.05$ corrected for multiple comparisons across the whole brain).

Fig. 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). Fig. 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 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). Fig.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), we have allowed more flexibility in the identification of map's borders following the methodology described in (Harvey et al., 2020).

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 highly 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.

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 $r_{run1,run2}$= 0.5, p= 0.11; $r_{run2,run3}$= 0.59, p = 0.04; $r_{run1,run3}$= 0.71, p = 0.007; long context $r_{run1,run2}$= 0.46, p = 0.04; $r_{run2,run3}$= 0.6, p= 0.005; $r_{run1,run3}$= 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). Finally, to exclude the possibility that the shift was driven by our winner-takes-all procedure and to the projection of the results on the cortical surface, we opted for a data-driven analysis on the volume space. We used cross-decoding (see Material and Methods) to compute the fraction of voxels able to learn the pattern of activity linked to the presentation of 0.65 s duration in the short context and to generalize it to the long context. Results showed (see S-Fig.19) that in SMA (as well as in IPS and parabelt) only a small fraction of voxels, 10.7% on average, displayed significantly above chance decoding accuracy.

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.

Fig.4AB shows the overlap between the maps of the two contexts (see Supplementary 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. Fig.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 was 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 (Protopapa et al., 2019), 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 Materials 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. Fig. 5 shows the presence of chronomaps in IPS (Fig.5A) and parabelt (Fig.5B) in a few individuals (for all individual maps see S-Fig.,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-Fig. 9 and 13 for individual maps), we observed very few vertices responding to 0.65 independently from the context (see also S-Fig. 19), and a spatial shift of vertices responding to 0.65 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-Fig. 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). Fig.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 both at whole brain level and ROI level (without considering the different durations, see S-Fig. 16, S-Fig. 18 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. We then wondered whether the similarities in the pattern of activity we found, could be used as proxy for understanding participants' behavior. To this aim we computed, for each participant, the dissimilarity matrix of their behavioral performance and then we correlated them with the dissimilarity matrix of the pattern of activity in SMA, IPS and parabelt at stimulus offset and reproduction onset (for comparisons, see Material and Methods). The result (see S-Fig. 20) showed that the pattern of brain activity at stimulus offset were associated with subject performance in all areas (median Pearson's r parabelt = 0.647, IPS = 0.657, SMA = 0.661) and the level of association was significantly higher ($F(1,45) = 17.01$ $p < 0.001$) as compare to response offset (median Pearson's r parabelt = 0.503, IPS = 0.429, SMA = 0.422).

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 (Harvey et al., 2020; Protopapa et al., 2019, 2020). 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 (Protopapa et al., 2019). 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 (Derdikman & Moser, 2010). 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 (Coull et al., 2015; Protopapa et al., 2019) and with the results of a recent effective connectivity study. 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 (Protopapa et al., 2020). 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 (Hayashi et al., 2015, 2018; Wiener et al., 2010). Both areas have been implicated in a variety of timing tasks (Bueti et al., 2008; Bueti & Macaluso, 2010; Wiener et al., 2010) with a range of durations spanning from a few hundreds of milliseconds to a few seconds (Lewis & Miall, 2003; Morillon et al., 2009) and with stimuli of different sensory modalities (Coull et al., 2008; Pastor et al., 2004). 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 (Kohn & Movshon, 2004). A change of response preferences measured with fMRI, has also been observed in numerosity maps after perceptual adaptation (Tsouli et al., 2021). 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 chronompas to other more “sensory” maps. Time maps together with numerosity maps have never been described in primary sensory cortices (Harvey et al., 2020; Protopapa et al., 2019), 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 (Zhou et al., 2018), 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 (Hendrikx et al., 2022). 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 (Lejeune & Wearden, 2009; Roach et al., 2017), only very few studies (Damsma et al., 2021; Murai & Yotsumoto, 2016) 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 (Damsma et al., 2021), 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., 0.65s) 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.

Materials And Methods

Participants

Fourteen healthy, right-handed volunteers (mean age 23 ± 3 years, mean ± 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^o 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 0.65s, 0.85s or 1.1s. 0.65s 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 (Kleiner et al., 2007) 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[33]. 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, TI_1/TI_2 =750/2350ms, α_1/α_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 (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. 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, 0.65s (SC), 0.65s (LC), 0.85s and 1.1s. 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.

Anatomical image processing

The high-resolution MP2RAGE images were analysed using Freesurfer software (Fischl et al., 2002) (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 (Destrieux et al., 2010).

Morphing using Freesurfer. Visualisations 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 in the area of interest and in a specific subject that was closest to the mean of that surface area 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 area closest to the mean of all the SMA label areas across subjects. This method of morphing ensures the best transformation of data from multiple sources to a single destination space.

Surface-based quantification of chronomaps spatial progressions

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

The areas where we explored the existence of chronomaps, 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.

Chronomaps 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 chronomaps 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:

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

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 0.65s 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 0.65s and reproduction of this duration across contexts, we plotted the difference in the reproduced duration of 0.65s in the two contexts (SC - LC) against the absolute difference in distance of the 0.65s 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^{o}, 90^{o} and 135^{o}. 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 visualise 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 visualised 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:

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

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 [6]). 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 (Oosterhof et al., 2016) 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 implementation: Chang & Lin, 2011) 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 visualised as confusion matrices (chance level is =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 ½=0.5. Using the same procedure we ran a whole-brain searchlight analysis at group-level. Firstly, the searchlight classification analysis (with spherical searchlights of 6 voxels of radius) was performed at the individual level, using beta values estimated from unsmoothed and MNI-aligned fMRI time-series (we excluded from this analysis 4 subjects that showed the highest degree of misalignment with the MNI template). Lastly, we computed the group-level statistics using *cosmo_montecarlo_cluster_stats* (using 10000 iterations and 'max' statistics).

Cross-decoding analysis. In order to check whether the spatial separations of activity clusters linked with the presentation of 0.65 s in the short and long context were not a byproduct of the analysis employed (see Surface-based quantification of chronomaps spatial progressions) we used a cross-decoding approach. This analysis was restricted to SMA, IPS and the auditory parabelt and to the dataset employed for the whole-brain searchlight analysis. This time since the analysis was carried on the volume we use JuBrain anatomy toolbox (Eickhoff et al., 2005) for ROI definition directly in the MNI space. The areas selected for the ROI definition were 6mc and 6mr for SMA, hIP1-8 for IPS and TE 1.2, TE 3 for the auditory parabelt. Cross-decoding was performed using searchlight analysis within each ROI (the radius of each spherical searchlight was set to 3 voxels). We trained a SVM classifier to discriminate

between the pattern of activity of 0.65 s in the short context and the activity of either 0.32 s or 1.1 s (this was done to avoid that the cross-decoding accuracy could be influenced by differences in beta values related to the different sessions). The classifier was then tested with the betas associated with 0.65 s in the long context. We used a cross-validation schema which included 18 independent combination of training and test set. For each voxel in each ROI we determined whether its accuracy was significantly above chance using a t-test against chance level (i.e. ½=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[38]. 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, Kriegeskorte, 2008; Kriegeskorte et al., 2006). 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 Pearson 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.

Additionally, we checked whether the RDM computed from the pattern of brain activity were correlated with subject behavior. To avoid that the different number of voxels per ROI in the different subjects could have an influence on this analysis we used the dataset employed for the group-level searchlight analysis and we use an absolute measure of similarity: the Euclidean distance. We first estimated the behavioral RDM computing the Euclidean distance between each pair of average reproduced durations in each participant using MATLAB's *pdist*. We then computed the RDM in all ROIs and participant, as described above, for both the betas associated with stimulus offset (stimulus RDMs) and response onset(response

RDMs). Finally, in each participant and ROI, we correlated behavioral RDMs with stimulus and response RDMs using Pearson's correlation coefficient, which were then Fisher transformed and analyzed using a linear mixed effect model with ROI and RDM type (either stimulus or response) as fixed effects and subject as random intercept. Group-level RDMs were summarized in S-Fig.20 using DISTATIS (Abdi et al., 2005).

References

Abdi, H., Valentin, D., O'Toole, A. J., & Edelman, B. (2005). Distatis: The analysis of multiple distance matrices. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2005-September*. https://doi.org/10.1109/CVPR.2005.445

Bueti, D., & Macaluso, E. (2010). Auditory temporal expectations modulate activity in visual cortex. *NeuroImage*, *51*(3), 1168–1183. https://doi.org/10.1016/j.neuroimage.2010.03.023

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience*, *20*(2), 204–214. https://doi.org/10.1162/jocn.2008.20017

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A Library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1–27. https://doi.org/10.1145/1961189.1961199

Coull, J. T., Charras, P., Donadieu, M., Droit-Volet, S., & Vidal, F. (2015). Sma selectively codes the active accumulation of temporal, not spatial, magnitude. *Journal of Cognitive Neuroscience*, *27*(11), 2281–2298. https://doi.org/10.1162/jocn_a_00854

Coull, J. T., Nazarian, B., & Vidal, F. (2008). Timing, storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. *Journal of Cognitive Neuroscience*, *20*(12), 2185–2197. https://doi.org/10.1162/jocn.2008.20153

Damsma, A., Schlichting, N., & van Rijn, H. (2021). Temporal context actively shapes EEG signatures of time perception. *Journal of Neuroscience*, *41*(20), 4514–4523. https://doi.org/10.1523/JNEUROSCI.0628-20.2021

Derdikman, D., & Moser, E. I. (2010). A manifold of spatial maps in the brain. In *Trends in Cognitive Sciences* (Vol. 14, Issue 12, pp. 561–569). Trends Cogn Sci. https://doi.org/10.1016/j.tics.2010.09.004

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, *53*(1), 1–15. https://doi.org/10.1016/j.neuroimage.2010.06.010

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data.

NeuroImage, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341–355. https://doi.org/10.1016/S0896-6273(02)00569-X

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: Applications. *NeuroImage*, *16*(2), 484–512. https://doi.org/10.1006/nimg.2002.1091

Glasauer, S., & Shi, Z. (2021). The origin of Vierordt's law: The experimental protocol matters. *PsyCh Journal*, *10*(5), 732–741. https://doi.org/10.1002/pchj.464

Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., & Paton, J. J. (2015). Striatal dynamics explain duration judgments. *ELife*, *4*(December2015). https://doi.org/10.7554/eLife.11386

Harvey, B. M., Dumoulin, S. O., Fracasso, A., & Paul, J. M. (2020). A Network of Topographic Maps in Human Association Cortex Hierarchically Transforms Visual Timing-Selective Responses. *Current Biology*, *30*(8), 1424-1434.e6. https://doi.org/10.1016/j.cub.2020.01.090

Hayashi, M. J., Ditye, T., Harada, T., Hashiguchi, M., Sadato, N., Carlson, S., Walsh, V., & Kanai, R. (2015). Time Adaptation Shows Duration Selectivity in the Human Parietal Cortex. *PLOS Biology*, *13*(9), e1002262. https://doi.org/10.1371/journal.pbio.1002262

Hayashi, M. J., van der Zwaag, W., Bueti, D., & Kanai, R. (2018). Representations of time in human frontoparietal cortex. *Communications Biology*, *1*(1), 1–10. https://doi.org/10.1038/s42003-018-0243-z

Hendrikx, E., Paul, J. M., van Ackooij, M., van der Stoep, N., & Harvey, B. M. (2022). Visual timing-tuned responses in human association cortices and response dynamics in early visual cortex. *Nature Communications*, *13*(1), 3952. https://doi.org/10.1038/s41467-022-31675-9

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

Kleiner, M., Brainard, D. H., Pelli, D. G., Broussard, C., Wolf, T., & Niehorster, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*.

Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience*, *7*(7), 764–772. https://doi.org/10.1038/nn1267

Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(NOV), 4. https://doi.org/10.3389/neuro.06.004.2008

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868. https://doi.org/10.1073/pnas.0600244103

Lejeune, H., & Wearden, J. H. (2009). Vierordt's the Experimental Study of the Time Sense (1868) and its legacy. In *European Journal of Cognitive Psychology* (Vol. 21, Issue 6, pp. 941–960). Taylor & Francis Group . https://doi.org/10.1080/09541440802453006

Lewis, P. A., & Miall, R. C. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia*, *41*(12), 1583–1592. https://doi.org/10.1016/S0028-3932(03)00118-0

Maaß, S. C., Schlichting, N., & van Rijn, H. (2019). Eliciting contextual temporal calibration: The effect of bottom-up and top-down information in reproduction tasks. *Acta Psychologica*, *199*, 102898. https://doi.org/10.1016/j.actpsy.2019.102898

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural Basis of the Perception and Estimation of Time. *Annual Review of Neuroscience*, *36*(1), 313–336. https://doi.org/10.1146/annurev-neuro-062012-170349

Merchant, H., Pérez, O., Zarco, W., & Gámez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *Journal of Neuroscience*, *33*(21), 9082–9096. https://doi.org/10.1523/JNEUROSCI.5513-12.2013

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience*, *12*(4), 502–507. https://doi.org/10.1038/nn.2272

Morillon, B., Kell, C. A., & Giraud, A. L. (2009). Three stages and four neural systems in time estimation. *Journal of Neuroscience*, *29*(47), 14803–14811. https://doi.org/10.1523/JNEUROSCI.3222-09.2009

Murai, Y., & Yotsumoto, Y. (2016). Context-Dependent Neural Modulations in the Perception of Duration. *Frontiers in Integrative Neuroscience*, *10*(MAR2016), 12. https://doi.org/10.3389/fnint.2016.00012

Oosterhof, N. N., Connolly, A. C., & Haxby, J. v. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*(JUL), 27. https://doi.org/10.3389/fninf.2016.00027

Pastor, M. A., Day, B. L., Macaluso, E., Friston, K. J., & Frackowiak, R. S. J. (2004). The Functional Neuroanatomy of Temporal Discrimination. *Journal of Neuroscience*, *24*(10), 2585–2591. https://doi.org/10.1523/JNEUROSCI.4210-03.2004

Paton, J. J., & Buonomano, D. v. (2018). The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. In *Neuron* (Vol. 98, Issue 4, pp. 687–705). Cell Press. https://doi.org/10.1016/j.neuron.2018.03.045

Protopapa, F., Hayashi, M. J., Kanai, R., & Bueti, D. (2020). Topographic Connectivity in a Duration Selective Cortico-Cerebellar Network. *BioRxiv*, 2020.04.08.031385. https://doi.org/10.1101/2020.04.08.031385

Protopapa, F., Hayashi, M. J., Kulashekhar, S., van der Zwaag, W., Battistella, G., Murray, M. M., Kanai, R., & Bueti, D. (2019). Chronotopic maps in human supplementary motor area. *PLOS Biology*, *17*(3), e3000026. https://doi.org/10.1371/journal.pbio.3000026

Roach, N. W., McGraw, P. v., Whitaker, D. J., & Heron, J. (2017). Generalization of prior information for rapid Bayesian time estimation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(2), 412–417. https://doi.org/10.1073/pnas.1610706114

Tsouli, A., Cai, Y., van Ackooij, M., Hofstetter, S., Harvey, B. M., te Pas, S. F., van der Smagt, M. J., & Dumoulin, S. O. (2021). Adaptation to visual numerosity changes neural numerosity selectivity. *NeuroImage*, *229*, 117794. https://doi.org/10.1016/j.neuroimage.2021.117794

van Ackooij, M., Paul, J. M., van der Zwaag, W., van der Stoep, N., & Harvey, B. M. (2022). Auditory timing-tuned neural responses in the human auditory cortices. *NeuroImage*, *258*, 119366. https://doi.org/10.1016/j.neuroimage.2022.119366

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage*, *49*(2), 1728–1740. https://doi.org/10.1016/j.neuroimage.2009.09.064

Zhou, J., Benson, N. C., Kay, K. N., & Winawer, J. (2018). Compressive temporal summation in human visual cortex. *Journal of Neuroscience*, *38*(3), 691–709. https://doi.org/10.1523/JNEUROSCI.1724-17.2017

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.

Figures

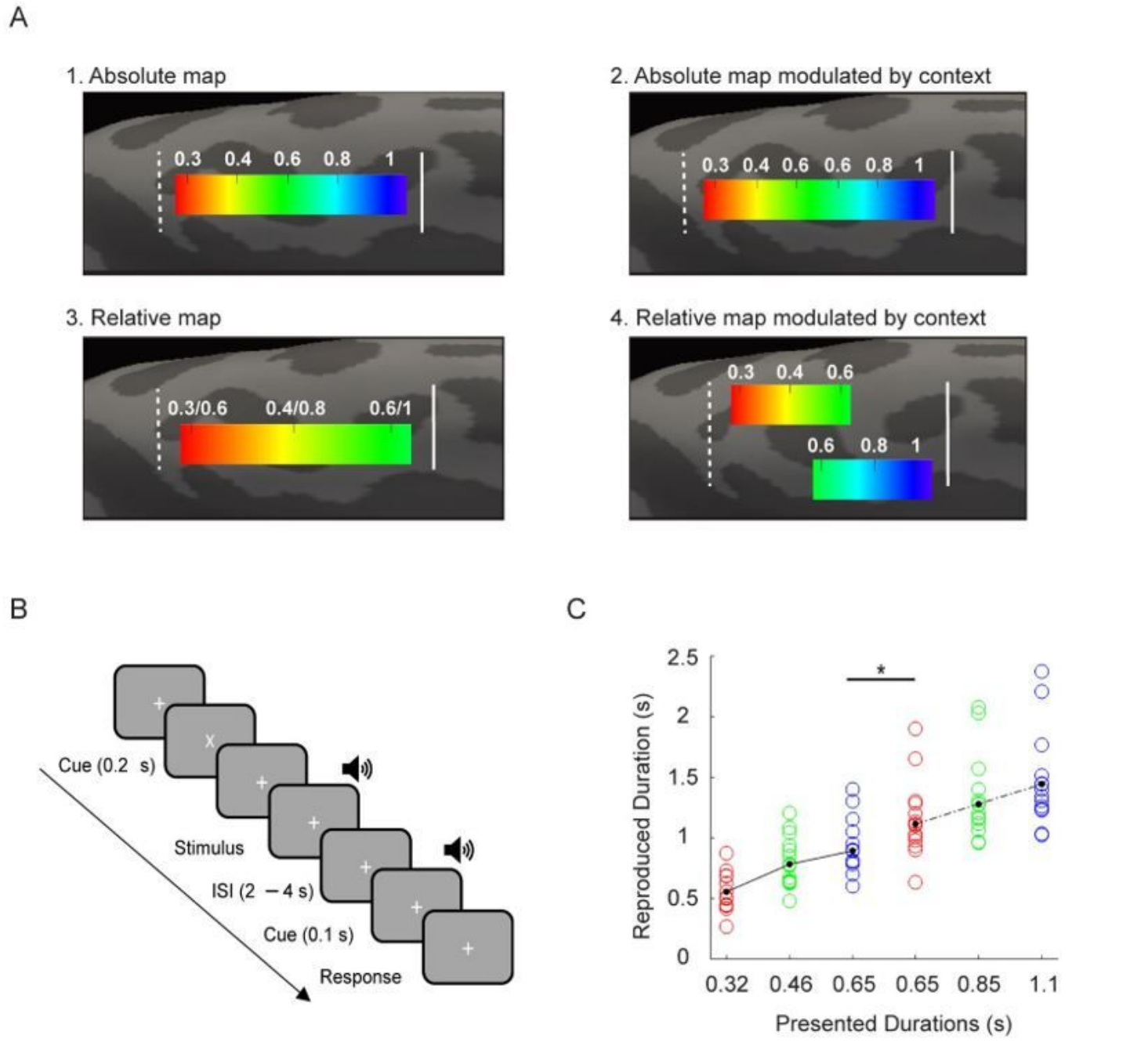

Figure 1

Predictions and experimental paradigm with behavioral results. (A) Sketch of four possible duration representations in chronomaps 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.

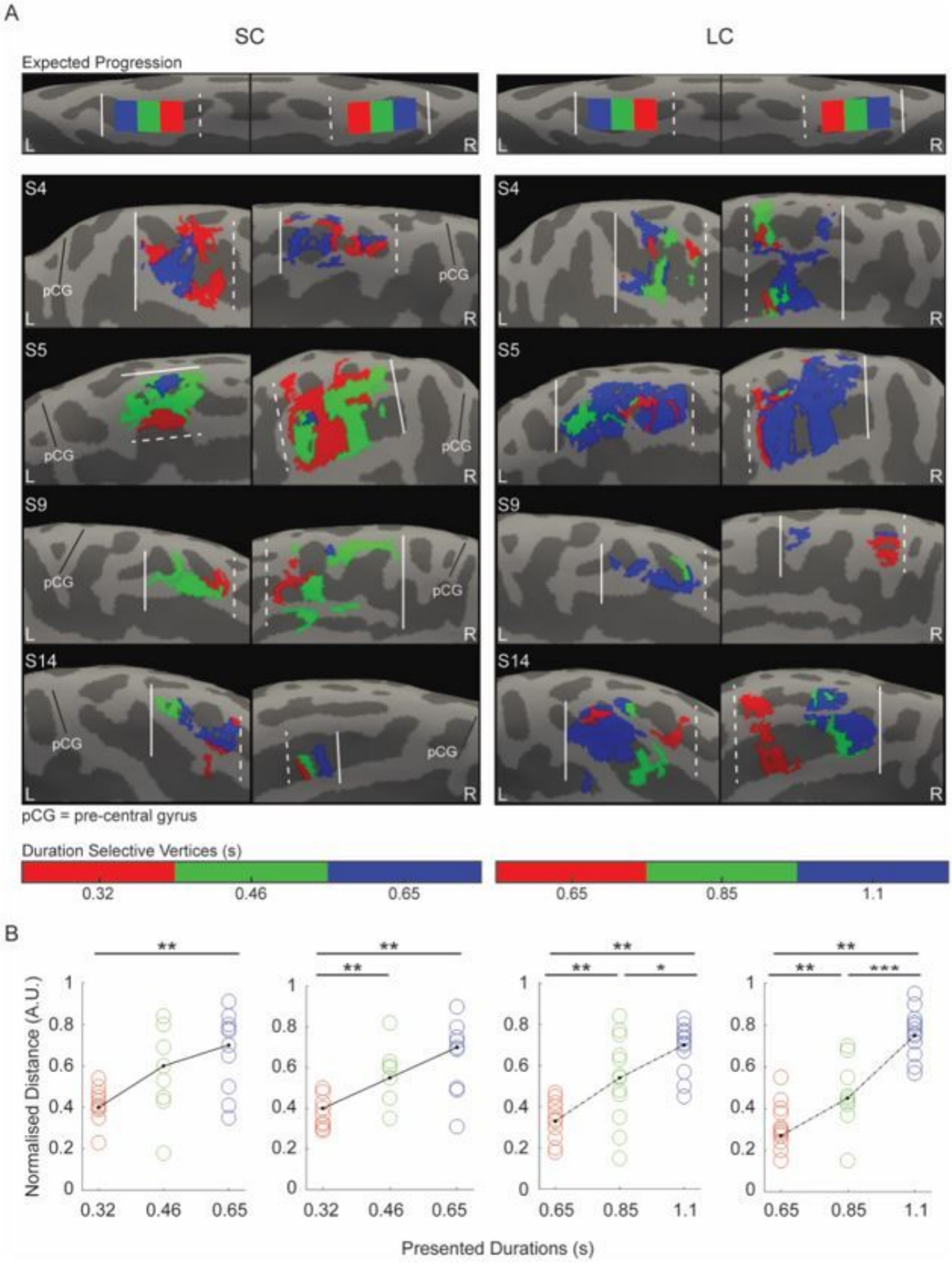

Figure 2

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. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

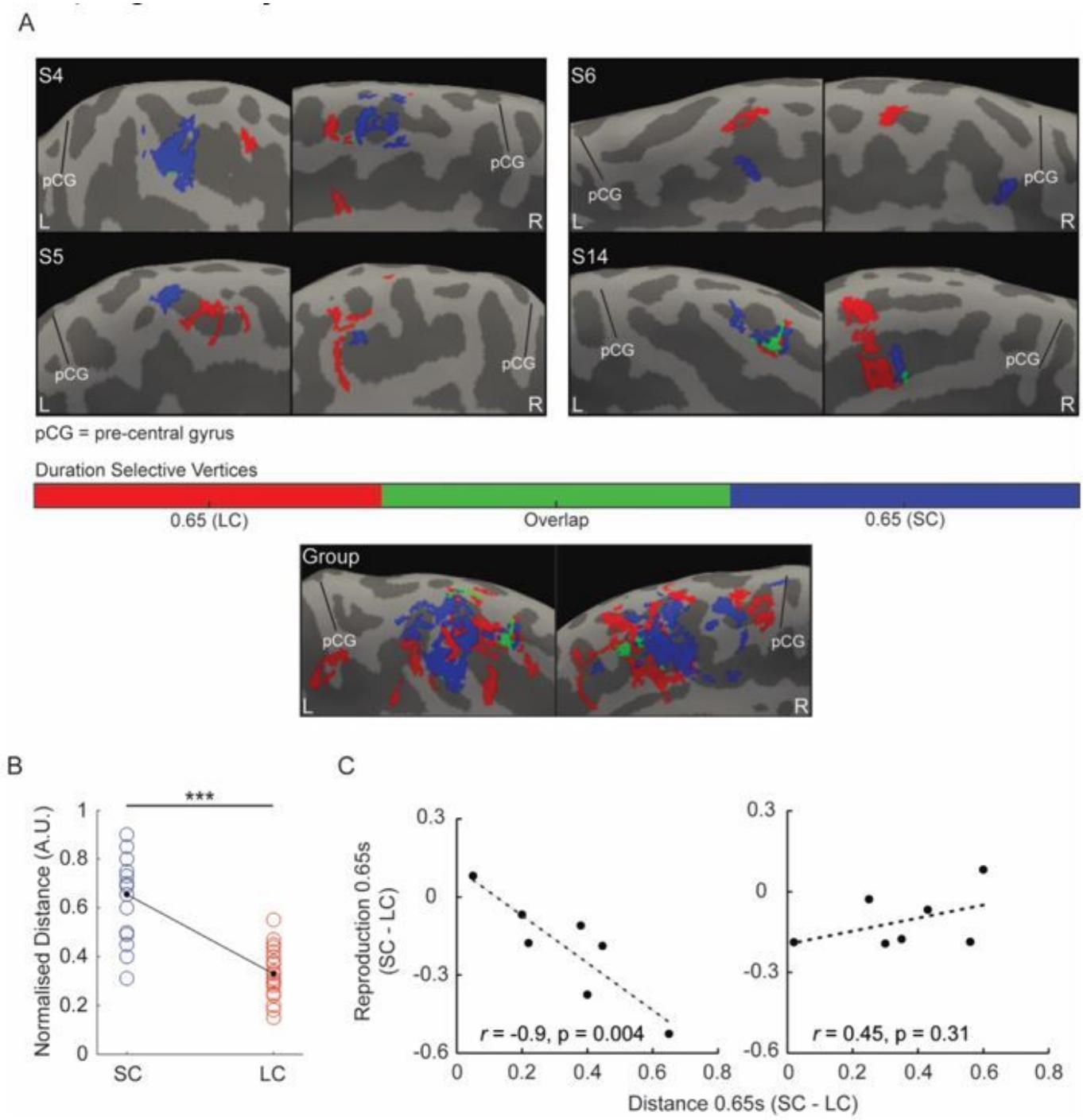

Figure 3

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 0.65s when presented in short context (SC, in blue) and long context (LC, in red). In green are the clusters responsive to 0.65 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(SC) = 15 and N(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 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 0.65s (y axis) and the normalized distance of the 0.65s 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).

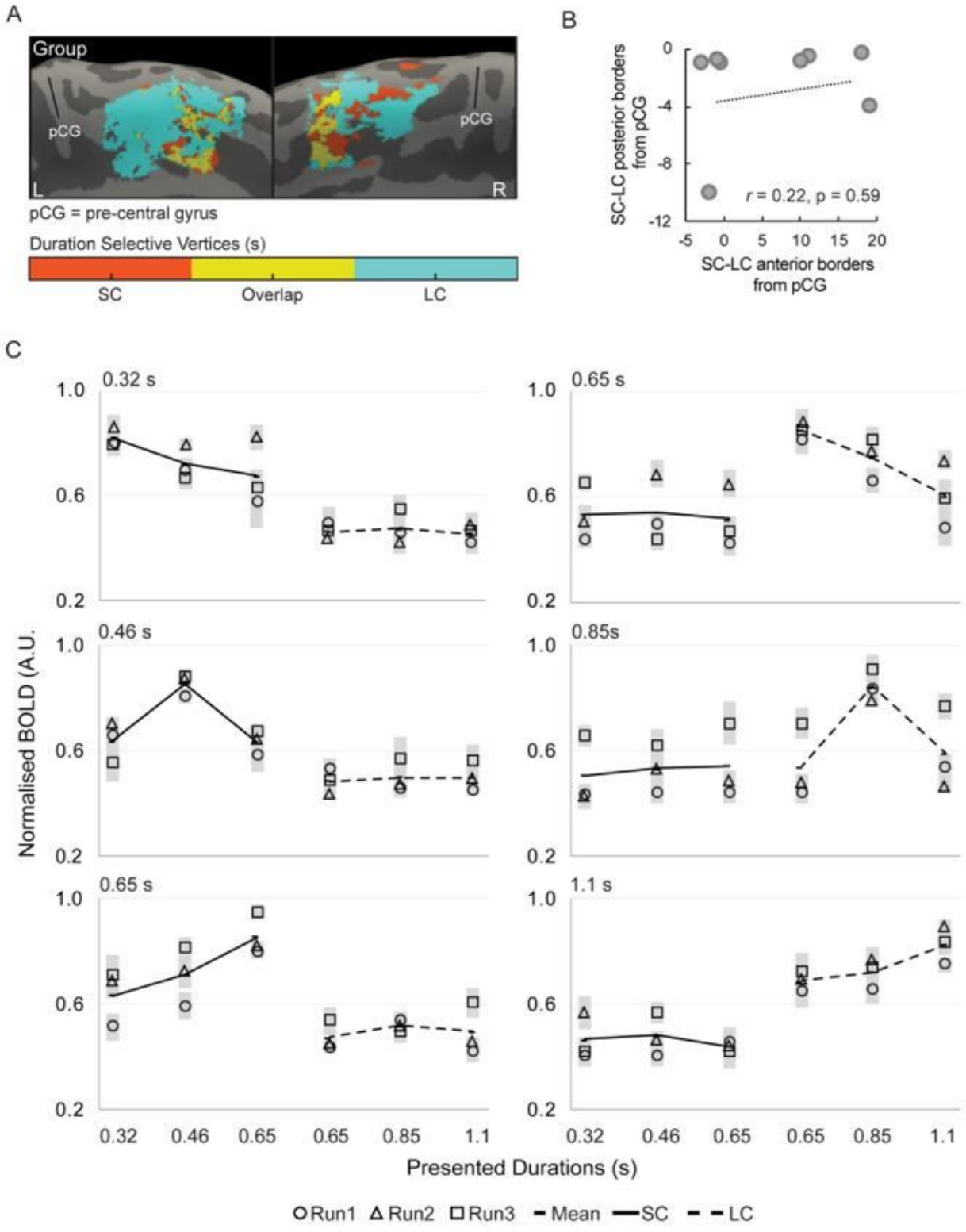

Figure 4

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 chronomaps where the progression was in the anterior to posterior orientation in the

two contexts in each hemisphere (N(left) = 5, N(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.

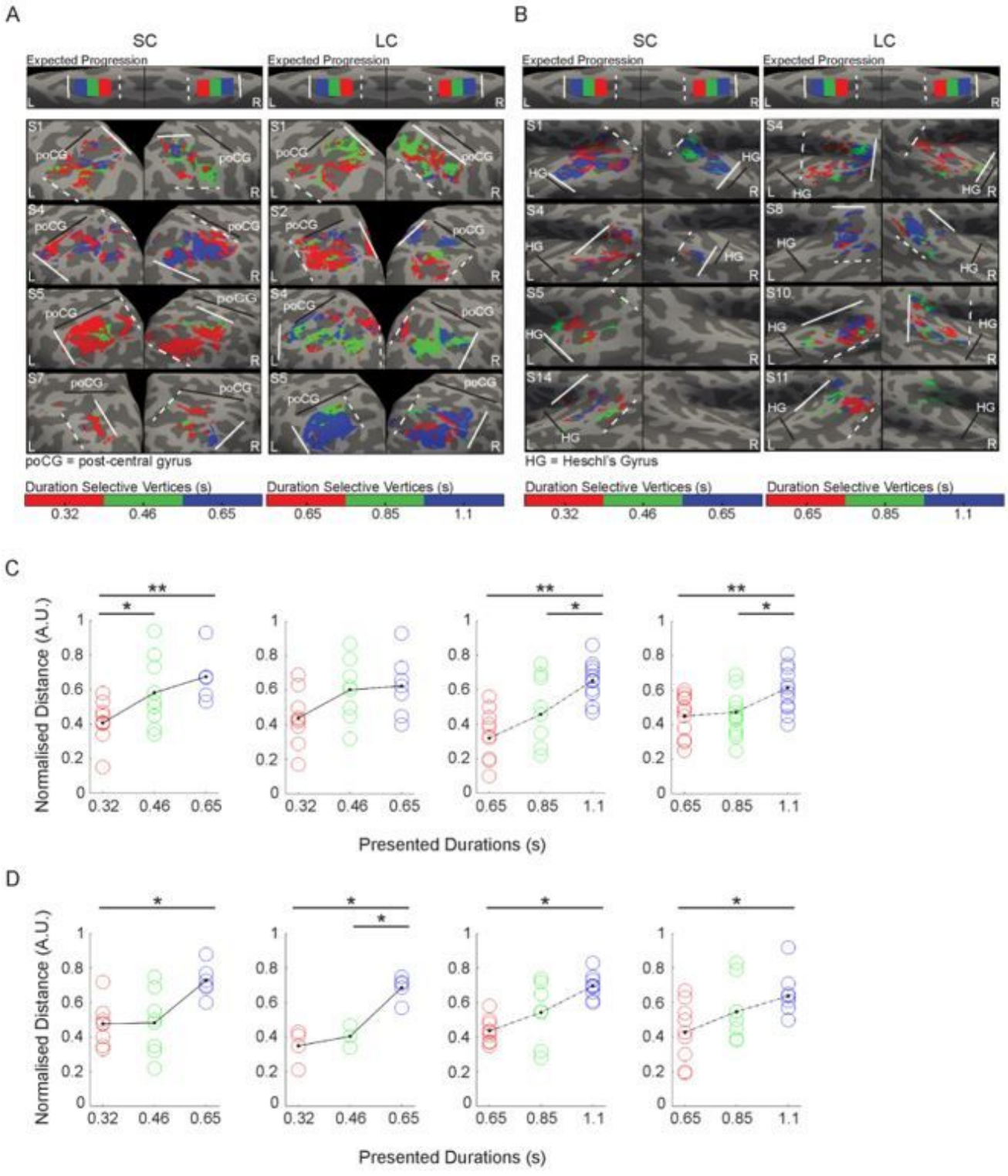

Figure 5

Auditory chronomaps 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. *$p < 0.05$, **$p < 0.01$

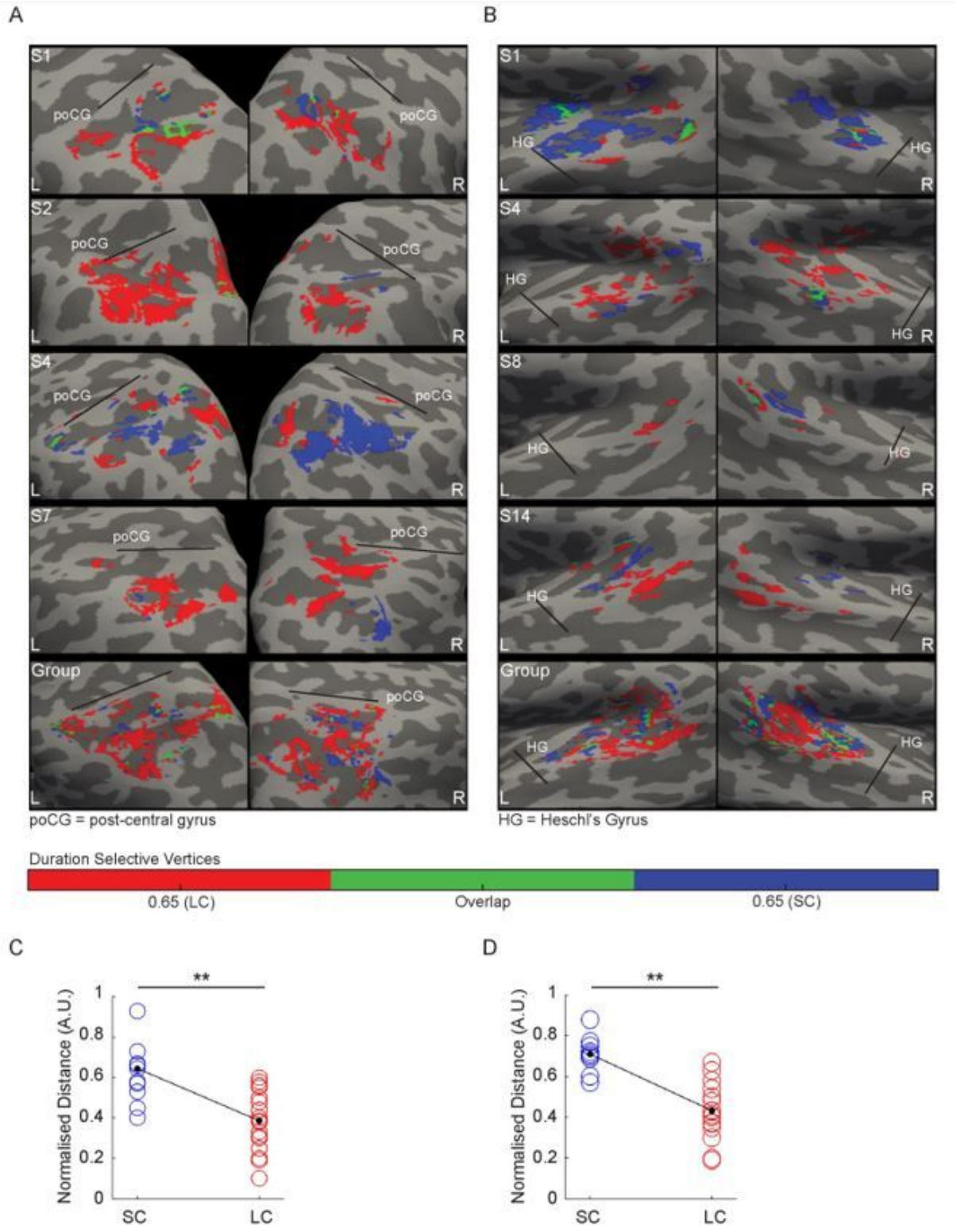

Figure 6

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 0.65s when presented in SC (blue) and LC (red). In green are the clusters responsive to 0.65 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 0.65s 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

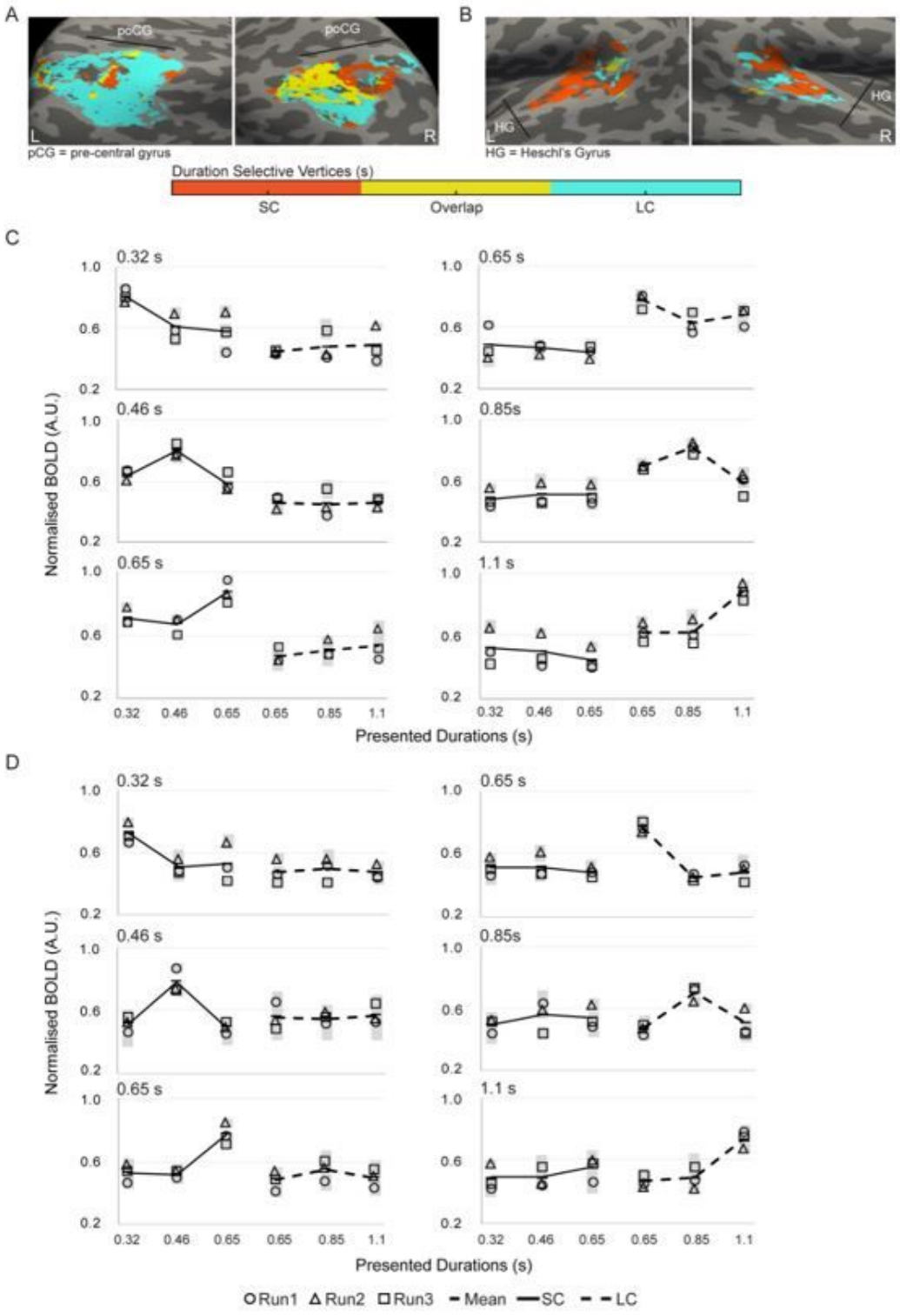

Figure 7

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 chronomaps where the progression was in the same

orientation in the two contexts and in each hemisphere (90o orientation in IPS N= 7 subjects and 180oorientation 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 normalised 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.

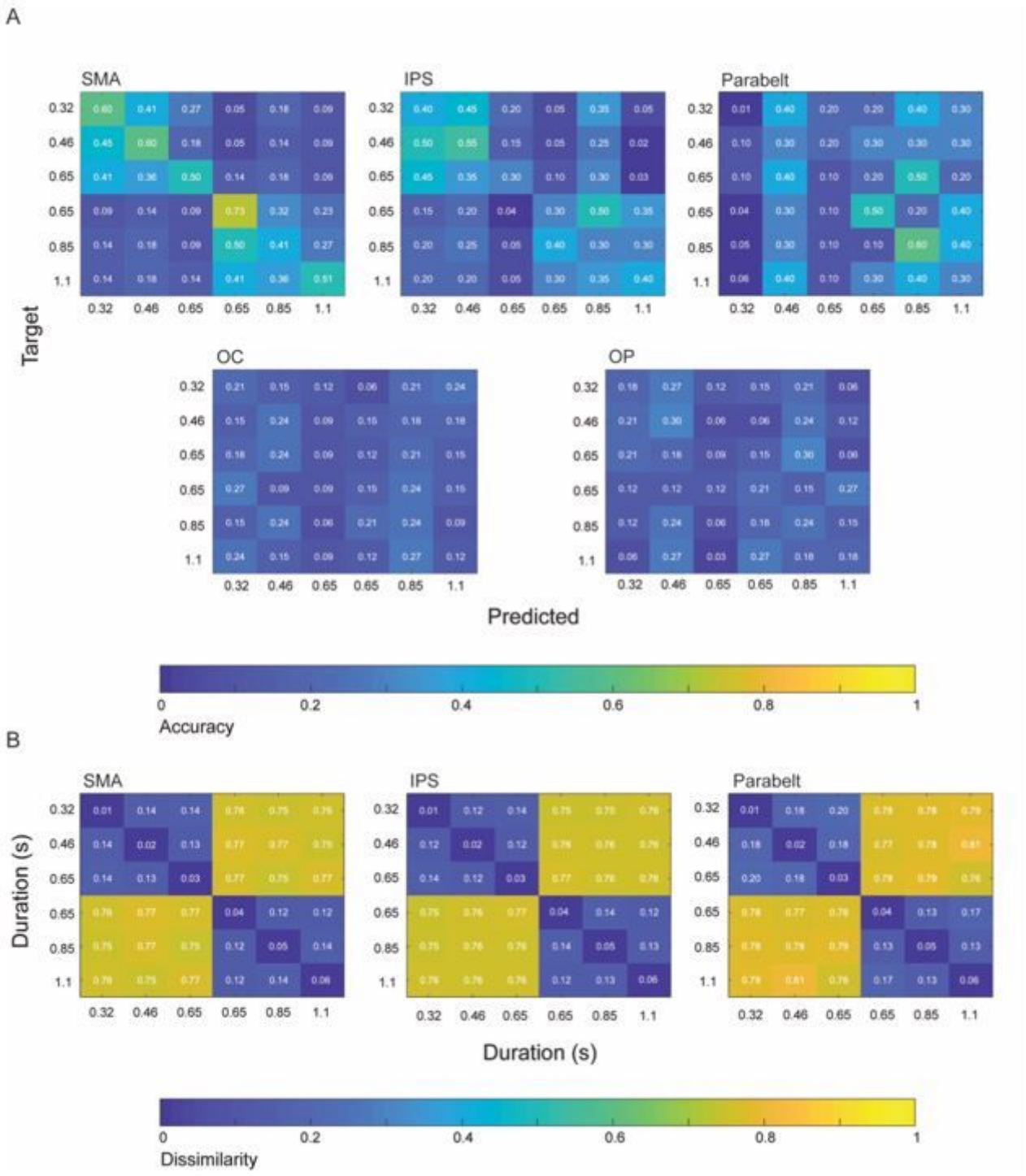

Figure 8

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). **(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 pare-wise (dis)similarity (1- Pearson's *r*) between each combination of the 6 β weights associated with the offset of the 6 durations The warmer the color, the greater the dissimilarity.

Supplementary Files

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

- SupplementaryMaterialKulashekharetal.v05.docx