Title

Prediction-related neural activity during silent periods is sharply tuned

Authors

Gianpaolo Demarchi*[1], Gaëtan Sanchez*[1,2] and Nathan Weisz[1]

Affiliations

1 - Centre for Cognitive Neuroscience and Division of Physiological Psychology, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria
2 - Lyon Neuroscience Research Center, CRNL, INSERM, U1028 – CNRS, UMR5292, Brain Dynamics and Cognition Team, Lyon, France

Author list footnotes

* shared first authorship

Corresponding author(s) e-mail address(es)

gianpaolo.demarchi@sbg.ac.at (G.D.)

Summary

Prior experience shapes sensory perception by enabling the formation of expectations with regards to the occurrence of upcoming sensory events. Especially in the visual modality, an increasing number of studies show that prediction-related neural signals carry feature-specific information about the stimulus. This is less established in the auditory modality, in particular without bottom-up signals driving neural activity. We studied whether auditory predictions are sharply tuned to even carry tonotopic specific information. For this purpose, we conducted a Magnetoencephalography (MEG) experiment in which participants passively listened to sound sequences of varying regularity (i.e. entropy). Importantly, sound presentation was occasionally omitted. This allowed us to assess whether and how carrier frequency specific information in the MEG signal is modulated according to the entropy level, especially during the silent (omission) periods. Using multivariate decoding analysis, our main finding is that only during an ordered (most predictable) sensory context does neural activity during omission periods contain carrier frequency specific information that can be used to classify neural activity elicited by genuine sounds. This shows that tonotopically specific patterns can be activated by top-down processes and supports the notion that predictions in the human auditory system can be sharply tuned.

Introduction

Our capacity to constantly predict incoming sensory inputs based on past experiences is fundamental to adapting our behavior in complex environments. A core enabling process is the identification of statistical regularities in sensory input, which does not require any voluntary allocation of processing resources (e.g. selective attention) and occurs more or less automatically in healthy brains (see e.g. (Ekman et al., 2017)). Analogous to other sensory modalities (Felleman and Van Essen, 1991; Ungerleider, 1994), auditory cortical information processing takes place in hierarchically organized streams along putative ventral and dorsal pathways (Rauschecker and Tian, 2000). These streams reciprocally connect different portions of auditory cortex with frontal and parietal regions (Plakke and Romanski, 2014; Rauschecker and Tian, 2000). This hierarchical anatomical architecture yields auditory cortical processing regions sensitive to top-down modulations, thereby enabling modulatory effects of predictions. In this context, a relevant question is to what extent do prediction-related top-down modulations (pre-)activate the same or similar neural ensembles as established for genuine sensory stimulation.

Such fine-tuning of neural activity would be suggested by frameworks that propose the existence of internal generative models (Friston, 2005; Rao and Ballard, 1999; Wolpert and Miall, 1996; Wolpert et al., 1995), inferring causal structure of sensory events in our environment and the sensory consequences of our actions. A relevant process of validating and optimizing these internal models is the prediction of incoming stimulus events, by influencing activity of corresponding neural ensembles in respective sensory areas. Deviations from these predictions putatively lead to (prediction) error signals, which are passed on in a bottom-up manner to adapt the internal model, thereby continuously improving predictions ((Friston, 2005); for an alternative predictive coding architecture see (Spratling, 2017)). According to this line of reasoning, predicted input should lead to weaker neural activation than input that was not predicted, which has been illustrated previously in the visual (Alink et al., 2010) and the auditory modality (Todorovic and de Lange, 2012). Support for the idea that predictions engage neurons specifically tuned to (expected) stimulus features has been more challenging to address and has come mainly from the visual modality (for review see (Summerfield and de Lange, 2014)). In an fMRI study Smith and Muckli (Smith and Muckli 2010) showed that early visual cortical regions (V1 and V2), which process occluded parts of a scene, carry sufficient information to decode above chance different visual scenes. Importantly, activity patterns in the occlusion condition are generalized to a non-occlusion control condition, implying context-related top-down feedback or input via lateral connections to modulate visual cortex in a feature specific manner. In a similar vein, it has been shown that mental replay of a visual stimulus sequence is accompanied by V1 activity that resembles activity patterns driven in a feedforward manner by the real sequence (Ekman, Kok, and de Lange 2017). Beyond more or less automatically generated predictions, explicit attentional focus to specific visual stimulus categories also goes along with similar feature-specific modifications in early and higher visual cortices even in the absence of visual stimulation (Peelen and Kastner 2011). Overall, for the visual

modality, these studies underline that top-down processes lead to sharper tuning of neural activity to contain more information about the predicted and / or attended stimulus (feature).

Studies as to whether predictions in the auditory domain (pre-)activate specific sensory representations in a sharply tuned manner are scarce especially in humans (for animal works see e.g. (Jaramillo and Zador, 2011; Rubin et al., 2016)). Sharpened tuning curves of neurons in A1 during selective auditory attention have been established in animal experiments ((Fritz et al., 2007a), even though this does not necessarily generalize to automatically formed predictions. A line of evidence could be drawn from research in marmoset monkeys, in which a reduction of auditory activity is seen during vocalization (Eliades and Wang, 2003). This effect is abolished when fed back vocal utterances are pitch shifted (Eliades and Wang, 2008). A recent work suggests that even inner speech may be sufficient to produce reduced neural activity, but only when the presented sounds matched those internally verbalized (Whitford et al., 2017). Using invasive recordings in a small set of human epilepsy patients, it was shown that masked speech is restored by specific activity patterns in bilateral auditory cortices (Leonard et al., 2016), an effect reminiscent of a report in the visual modality ((Ekman et al., 2017); for other studies investigating similar auditory continuity illusion phenomena see (Kraemer et al., 2005; Müller et al., 2013; Voisin et al., 2006)). Albeit being feature specific, this “filling-in” type of activity pattern observed during phoneme restoration cannot clarify conclusively whether they require top-down input. In principle these results could also be largely generated via bottom-up thalamocortical input driving feature relevant neural ensembles via lateral or feedforward connections. To resolve this issue, putative sharp tuning via predictions needs to be shown absent of feedforward input (i.e. silence). Furthermore, the exact timing of the effects could provide important evidence on whether predictions come along with feature-specific preactivations of relevant neural ensembles (Summerfield and de Lange, 2014).

The goal of the present study was to investigate in healthy human participants whether predictions in the auditory modality are exerted in a carrier frequency (i.e. tonotopic) specific manner. For this purpose, we merged an omission paradigm with a regularity modulation paradigm (for overview see (Heilbron and Chait, 2017), and see **Figure 1** for the specific details). So-called omission responses occur when an expected tone is replaced by silence. Frequently this response has been investigated in the context of Mismatch Negativity (MMN; (Näätänen et al., 2007)) paradigms, which undoubtedly have been the most common approach of studying the processing of statistical regularities in human auditory processing (e.g. (Bendixen et al., 2014; Recasens and Uhlhaas, 2017; Sanmiguel et al., 2013)). This evoked response occurs upon a deviance from a “standard” stimulus sequence, that is, a sequence characterized by a rule endowing it with a certain degree of order. For omission responses (e.g. (Raij et al. 1997)), this order is usually established in a temporal sense, that is, allowing precise predictions when a tone will occur ((Wacongne et al., 2011); for a study using a repetition suppression design see (Todorovic et al., 2011)). The neural responses during these silent periods are of outstanding interest since they cannot be explained by any feedforward propagation of activity elicited by a physical stimulus. Thus, omission of an acoustic stimulation will lead to a neural response, as long as this omission violates a regular sequence of acoustic stimuli, that is, it occurs unexpectedly. Previous works have identified auditory cortical contributions to the omission response (e.g. (Todorovic et al., 2011)).

Interestingly, and underlining the importance of a top-down input driving the omission response, a recent DCM study by Chennu et al. (Chennu et al., 2016) illustrates that it can be best explained when assuming top-down driving inputs into higher order cortical areas (e.g. frontal cortex). While establishing temporal predictions via a constant stimulation rate, we varied the regularity of the sound sequence by parametrically modulating its entropy level (see e.g. (Auksztulewicz et al., 2017; Barascud et al., 2016)). Using different carrier frequencies, sound sequences varied between random (high entropy; transition probabilities from one sound to all others at chance level) and ordered (low entropy; transition probability from one sound to another one above chance). Our reasoning was that omission-related neural responses should contain carrier frequency specific information that is modulated by the entropy level of the contextual sound sequence. Using a time generalization decoding approach (King and Dehaene, 2014), we find evidence that particularly during the low entropy (highly ordered) sequence, neural activity in the omission period contains carrier frequency specific information similar to activity observed during real sound presentation. This work shows that prediction-related neural activity in the auditory system are sharply tuned even down to the tonotopic level.

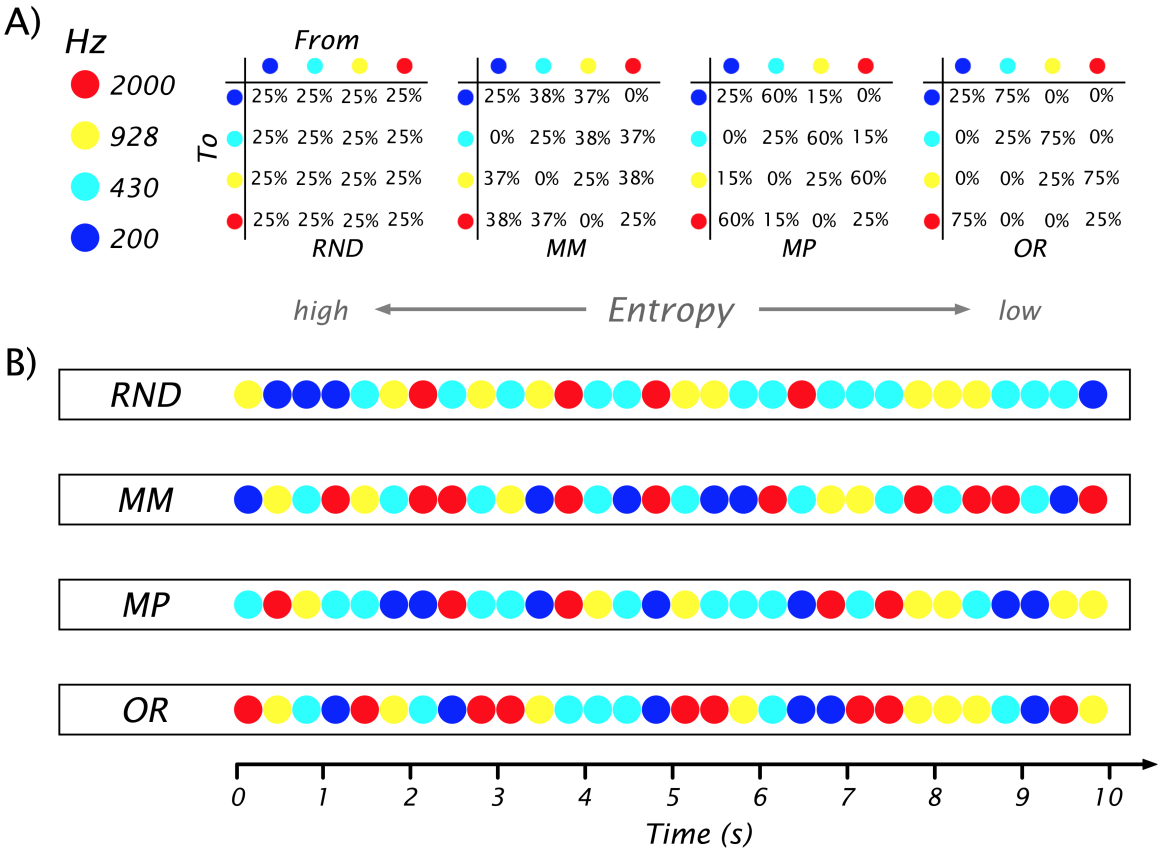

Figure 1

Experimental design. A) Transition matrices used to generate sound sequences according to the different conditions (RND, MM, MP and OR). **B)** Schematic examples of different sound sequences generated across time. 10% of sound stimuli were randomly replaced by omission trials (absence of sound) in each context.

Results

Sound and omission evoked responses show differential relationship with entropy

We found clear sound- and omission-related cortical evoked responses at the grand average level. Disregarding the specific entropy level as well as carrier frequency, during a time period of 50-200 ms post event onset, striking overlaps between sound and omission evoked generators were observed in the right primary auditory cortex (A1; see **Figure 2A**, left panel). Despite this spatial overlap, the temporal dynamics in right A1 showed a faster peak when an actual sound was presented as compared to when an omission occurred (see **Figure 2A**, right panel). Outside of right A1, pronounced evoked responses were also observed for sounds and omissions in an idiosyncratic manner: while left A1 was also strongly activated by sounds, omissions went along with strong evoked activity in the primary visual cortex (V1; see **Figure 2A**, left panel). The latter may be associated with the fact that unexpected omissions may involuntarily lead to an orientation response, which involves visual exploration. Since this issue was not relevant to the research focus it was not further followed up.

To assess the relationship between evoked activity and the entropy level, we performed a regression analysis in time-windows 100-200 ms and 200-300 ms post-event onset. For the early (100-200 ms) time-window, no effect at the cluster corrected level was observed (all p 's > .1). In the later (200-300 ms) time-window, lower entropy (more regularity) was reflected in weaker evoked responses (negative cluster: p = .007). This effect showed maxima at 263 and 297 ms after sound onset and was localized to precuneus and right striatum for both time points (**Figure 2B**, left panel). Right A1 was also implicated, but only for the later period (**Figure 2B**, left and right panel). No effect at a cluster-corrected level was obtained for omissions (p 's > .19). Given our hypothesis that contrary to sound evoked responses, omission evoked responses should increase with decreasing entropy, we applied a regression test to a normalized contrast restricted to left and right A1 (**Figure 2A**). This measure normalized each condition with respect to the highest entropy (random) condition and the difference between sound and omission within each entropy level was entered into the statistic. For the left A1 a significant cluster corrected effect was obtained at later time points (250-300 ms; right A1: n.s.; see **Figure 2C**, left panel), which was driven by a differential relationship with entropy level for sounds and omissions.

Overall, the analysis of evoked responses establishes common generators for omissions and genuine sounds in right A1. This region and a set of non-auditory regions (precuneus and right striatum) show decreasing sound evoked responses with increasing regularity of the sound sequence. Increasing omission evoked activity was observed with increasing regularity of the sound sequence in left A1.

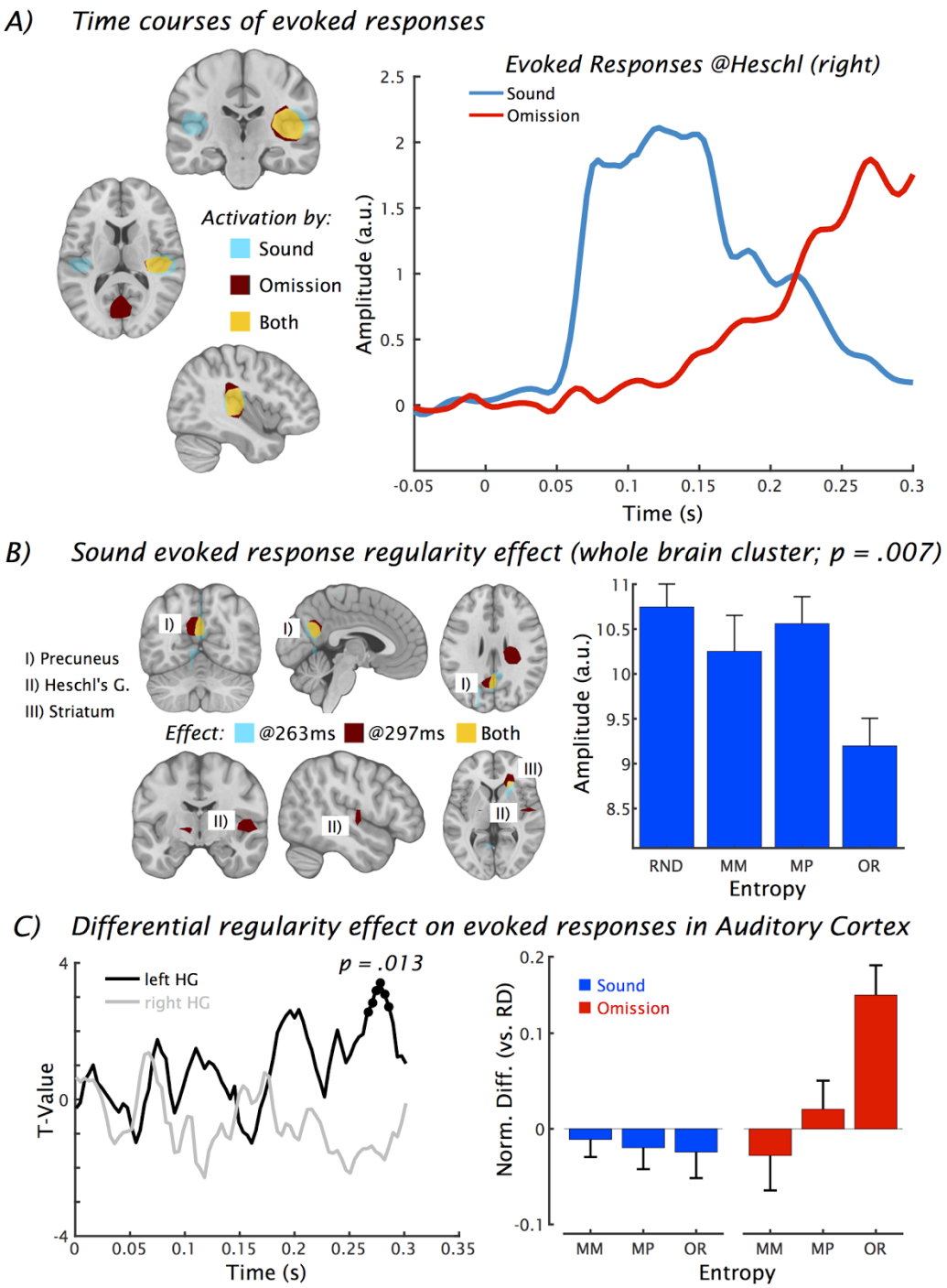

Figure 2

Sound and omission evoked response modulation according to entropy level. (A) Source localization of evoked brain activity in response to sound or omission between 50 and 200 ms (left panel) and time-course of source activity in right Heschl's gyrus (right panel). Brain source clusters are masked at 85% maximum activity. All trials were averaged separately for sounds and omission disregarding entropy level and tone frequency. For visualization purposes, a baseline between -50 ms to sound / omission onset was subtracted to the averaged time-course depiction. **(B)** Linear regression analysis over different contextual regularity conditions (4 entropy levels) for time period between 200 and 300 ms

Single-trial neural activity during sound contains information about entropy level

The results in the previous step were obtained after trial averaging, which leaves the question open as to what information is contained in the signal on a single trial level. To validate our decoding approach, we first followed up on the strong evoked effect showing differential sound evoked response amplitudes for the different entropy levels. Using all magnetometers (i.e. discarding the spatial pattern) and a time-generalization decoding approach showed that the entropy level of the condition in which the sound was embedded into could be decoded above chance from virtually any time point and generalized to any other time point. Only the on-diagonal result is depicted in **Figure 3** (left panel; the off-diagonal patterns will be part of a separate manuscript), showing globally above chance level decoding accuracy. This finding of temporally generalizable non-specific neural patterns fits well with the fact that conditions were presented in blocks. A transient increase following ~100-200 ms post-stimulus onset can be observed, which is somewhat earlier than the evoked response effect described above. This underlines that the outcome of this decoding analysis is not merely an alternative depiction of the evoked response analysis. In order to identify potential neural generators that drive the described effect, the time-generalization analysis was followed up by a searchlight analysis in source space. Since sensor level analysis suggested a temporally stable (in the sense of almost always significant) neural pattern, the entire 0-300 ms time period was used for this purpose. The analysis revealed above chance decoding accuracy spread throughout almost the entire brain. In order to identify potential “hot-spots” a 10% of maximum decoding accuracy threshold was introduced, showing that the largest effect was obtained in the right hemisphere, encompassing large portions of the temporal cortex. Based on this analysis, we can state that information about the regularity of the sound sequence is contained also at the single-trial level in a temporally stable manner and the right temporal cortex may play a pronounced role in representing the regularity of sound sequences.

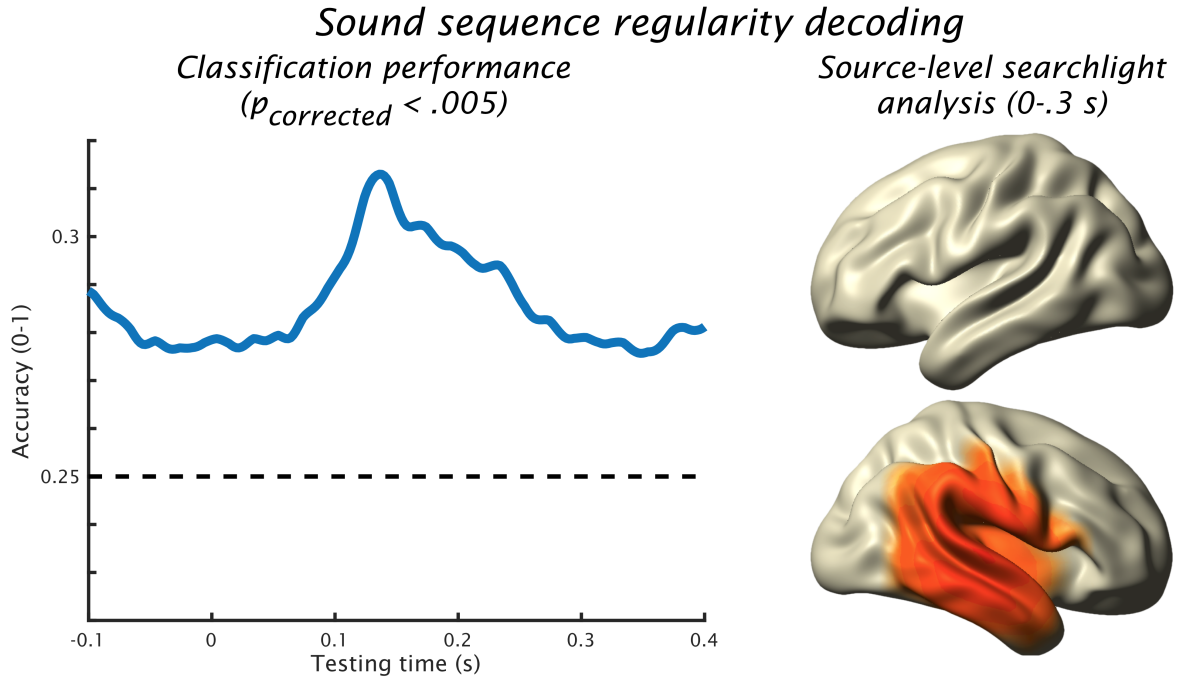

Figure 3

Decoding of entropy level. Outcome of decoding accuracy of the on-diagonal shows temporally stable above chance scores, with a transient increase following ~100-200 ms post-stimulus onset (left panel). Searchlight analysis on source level using a wide temporal window (0-300 ms) shows a widely distributed pattern with a maximum in right temporal regions (image thresholded at 10% of maximum).

Single-trial neural activity during sound contains information about tone frequency

Prior to addressing the more challenging question of whether silent periods (i.e. omissions) contain carrier frequency specific information (omission-to-sound decoding), we first tested whether this was in general possible using neural activity to actual sound presentation (sound-to-sound decoding). In an analogous approach to the one previously described, we derived time-generalized carrier frequency decoding performance separately for the different entropy levels. Sound frequency could be decoded high above chance from MEG activity, disregarding the regularity of the sound sequence (**Figure 4A**). For all entropy levels, a temporally relatively stable pattern emerges between 100 to 300 ms, with the highest accuracy clustering along the diagonal. The off-diagonal pattern becomes descriptively more pronounced with increasing regularity of the sound sequence. This impression is confirmed by statistically testing a linear trend (**Figure 4B**; left) showing a significant neural pattern at ~200 ms training time generalizing for ~100 ms. This effect is observed in spite of sound evoked responses showing overall decreased amplitudes with increasing regularity of the sequence (**Figure 2B**). We investigated potential generators for this significant neural pattern with a source-level searchlight decoding analysis performed within this significant

time window from the time-generalization matrix. This analysis shows a large pattern involving mainly left superior temporal regions, left inferior frontal and left inferior and superior parietal regions. Furthermore, occipital and deep regions are also visible such as: left middle occipital gyrus, left calcarine sulcus, precuneus and anterior cingulate gyrus.

Prior to this effect, a further pattern of increasing accuracy emerges as a function of entropy levels between ~-100 to ~120 ms, yielding a significant linear trend (**Figure 4B**; left). This pattern consists of a strong on-diagonal part, monotonously decreasing in strength as a function of temporal distance. In principle, it could contain some prediction-related preactivations. However, given the setup of the experiment, the effect is likely generated in large part by carry-over neural activity of the previous tone, that is exploited by the classifier in decoding the present tone frequency. However, another pattern within this time frame is an increasing off-diagonal decoding accuracy emerging ~80-100 ms post-stimulus and extending from pre-sound periods to almost 200 ms (see **Figure 4B**; left). The time almost coincides with the overall onset on-diagonal increase in decoding accuracy, but the 45° orientation of this effect with respect to the diagonal makes it unlikely that it is artifactual. The pattern for the pre-sound period indicates that while carry-over is certainly a part, some parts are similar to neural patterns driven by the new sound. To test this impression more formally, we averaged decoding accuracy within selected slices from the time generalization matrix that reflected how strongly a pre-sound neural pattern generalizes over time (see **Figure 4C**, left). Assuming that the influence of the preceding sound (and thereby the carry-over) should be strongest at earliest intervals and subsequently become weaker, temporal decoding accuracy profiles between -70 and 150 ms were modeled by linear regression (see **Figure 4C**, middle). For every time point, deviations from the carry-over model were assessed by calculating the residuals from the linear fit for every time point in this window. Within the aforementioned time-window of 80-100 ms, residuals appeared to increase as a function of the entropy level. Statistically significant deviations were only obtained for non-random sound sequences, being particularly pronounced in the ordered condition at 80 to 90 ms. Altogether the results from this analysis are difficult to reconcile with carry-over neural activity from the previous tone (i.e. neural patterns elicited by a tone ~80-100 ms being the same as the pattern to a previous tone of a different frequency), but instead are more parsimoniously explained by test-tone carrier-frequency specific effects that are already present prior to actual stimulus onset. This would speak in favor not only of the predictability of the sound sequence to affect sound processing in tonotopically specific manners, but also that this manipulation could instantiate tonotopically specific preactivations.

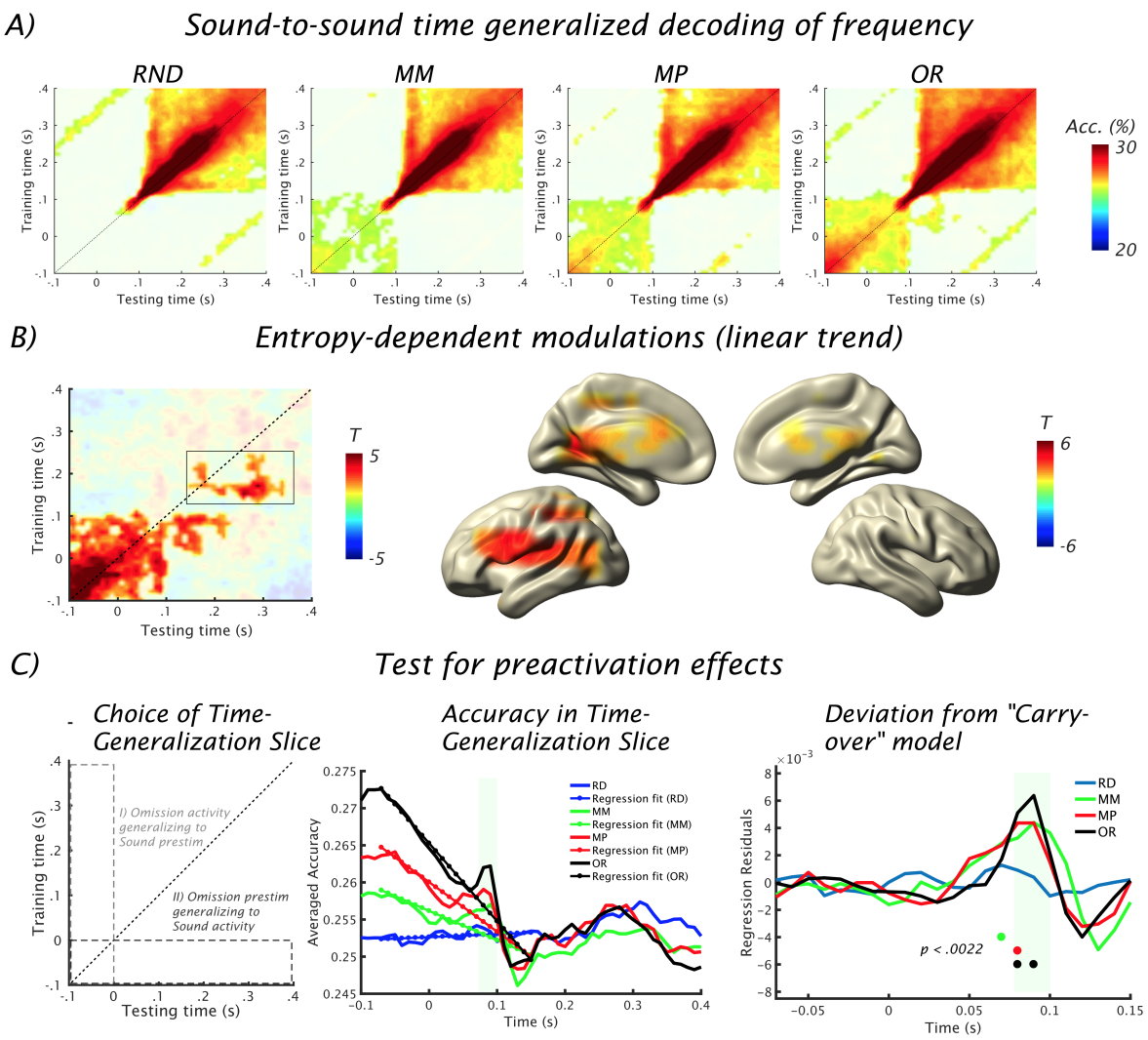

Figure 4

Sound-to-sound decoding results. (A) Time-generalization analysis on sensor data (magnetometers) for each entropy level. Time-by-time decoding was performed on data restricted to sound trials. Classifiers were trained at different time points to decode tone-frequency (4 classes). Each matrix presents the result of decoding for each entropy level where classifier accuracy was significantly above chance level (25%) masked at $p_{corrected}$<0.005. **(B)** Linear trend analysis between entropy levels. On the left panel, the time generalization matrix represents significant linear trend for decoding accuracy values over the group (masked at $p_{corrected}$<0.05). The significant time-by-time points inside the black square depiction are used to investigate source level decoding searchlight linear-trend analysis on right panel (masked at $p_{corrected}$<0.05). **(C)** Pre-activation pattern analysis for each entropy level. Training time points from -100 ms to 0 ms were averaged over the testing time course (see dashed black rectangle in left panel). A regression linear fit was used to remove previous stimulation carry-over activity (middle panel). One can observe a residual accuracy increase around 90 ms in testing time for all context but random (RND) speaking in favor of pre-stimulation activation patterns related to tone-frequency prediction.

Frequency of expected but omitted tones can be decoded only during regular sound sequences

Our analyses so far show that regularity of the sound sequence affects neural responses to sounds and even influences the performance of a classifier to decode tone frequencies. However the putative effects of predictions up to this point (except of the omission evoked response; see **Figure 1C**) have been obtained in the presence of a sound. This is not sufficient to address our main question of whether prediction-mediated neural processes are of sufficient granularity to contain information also about what sound was predicted in absence of any acoustic information. Pursuing an analogous time-generalized decoding approach as previously described, we tested whether neural patterns around omissions (training set) can be found during genuine tones (test set). Indeed activity ~150-250 ms following onset of the omission could classify significantly above chance level sound frequency in the test data set. However, this was only the case for the ordered condition (**Figure 5A**).

Testing a linear trend across the time-generalization decoding results confirms this general pattern (**Figure 5B**, left). Using a searchlight analysis on source level data (thresholded at $p_{corrected} < .01$), we followed up on probable generators of the sensor level effect. For this purpose we focused on the late significant effect, that is, between ~100 and 200 ms training time (from omissions) tested on neural patterns ~250 ms following sound onset. This analysis shows a right hemispheric dominant pattern encompassing particular regions of the auditory cortex, but also motor and premotor regions (**Figure 5B**, right). Furthermore, a strong linear trend was also identified in anterior cingulate cortex and subcortical structures such as hippocampus and thalamus.

Also, at the 90 ms testing time period, the activity bears similarities to the one recorded in the pre-omission period. This effect is reminiscent of the one reported above (see **Figure 4**) and we followed up on the question of whether this is a trivial carry-over effect or one indicating a test-tone frequency specific preactivation effect using the same approach. However, here we used only the time-generalization slice that contains information as to whether pre-omission neural patterns can be found during tone presentation. As previously reported, deviations from a linear regression fit were most pronounced at 90 ms (see **Figure 5C**, middle), which is significant albeit at a non-Bonferroni-corrected level (see **Figure 5C**, right). In light of the more strongly powered analysis in the previous section, we take this as corroborating evidence that neural patterns in the frequency of the test tone are already present prior to the presentation of the actual sound underlining the proactive nature of predictions. Most importantly, however, the results in this section unequivocally show that regularity of the sound sequence - putatively modulating predictions - lead to tonotopically neural activity patterns during omission periods.

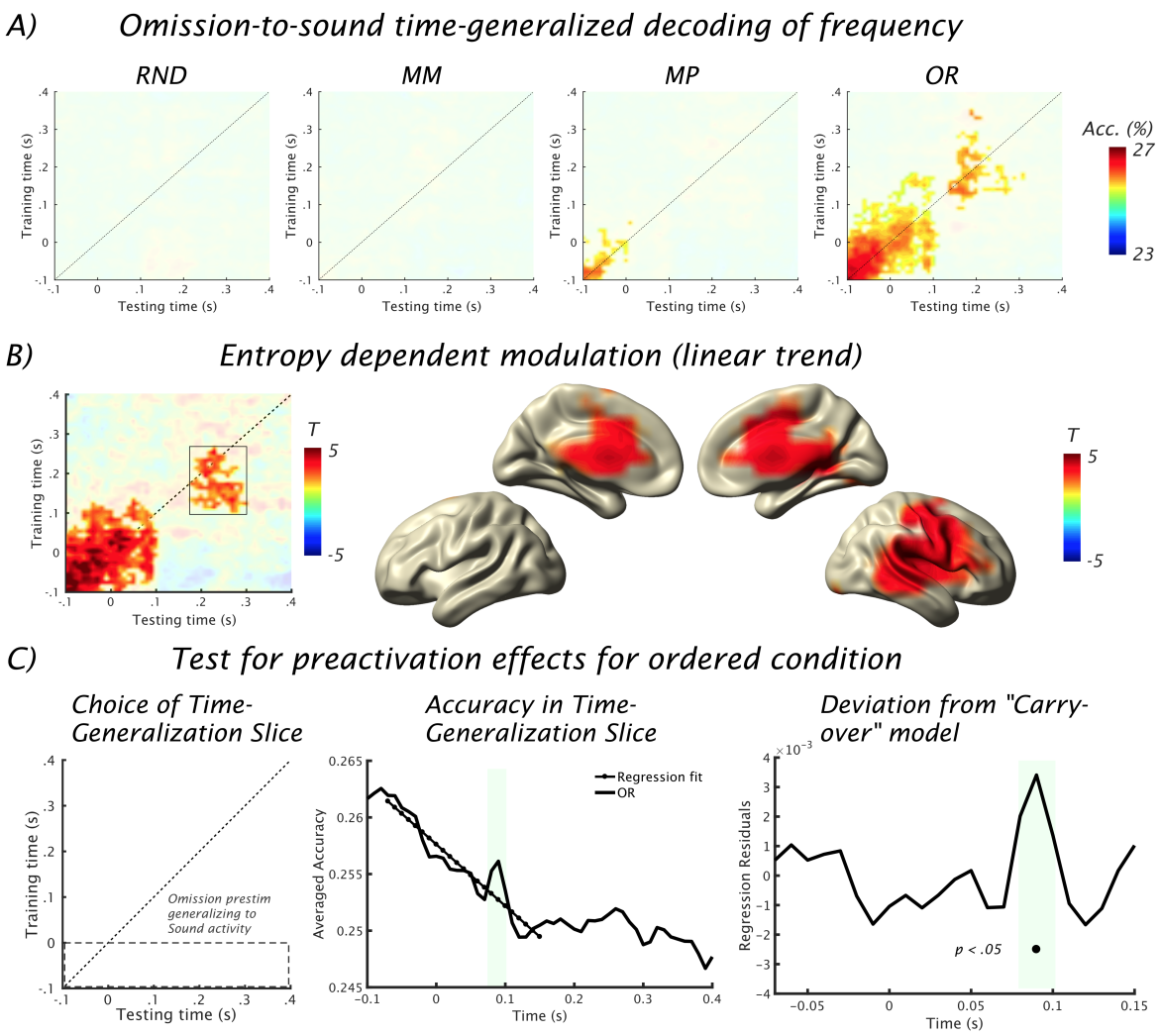

Figure 5

Omission-to-sound decoding results. (A) Time-generalization analysis on sensor data (magnetometers) for each entropy level. Omission trials were used as training data and sound trials as testing data. Classifiers were trained at different time points to decode tone-frequency (4 classes). Each matrix presents the result of decoding for each entropy level where classifier accuracy was significantly above chance level (25%) masked at $p_{corrected}$<0.005. Only MP and OR entropy conditions show significant decoding at the group level. **(B)** Linear trend analysis between entropy levels. In the left panel, the time generalization matrix represents a significant linear trend for decoding accuracy values over the group (masked at $p_{corrected}$<0.05). The significant time-by-time points inside the black square depiction are used to investigate source level decoding searchlight linear-trend analysis on the right panel (masked at $p_{corrected}$<0.01). **(C)** Pre-activation pattern analysis for ordered context. Training time points from -100 ms to 0 ms were averaged over the testing time course (see dashed black rectangle in left panel). A regression linear fit was used to remove previous stimulation carry-over activity (middle panel). One can observe a residual accuracy increase around 90 ms in testing time speaking in favor of pre-stimulation activation patterns related to tone-frequency prediction.

Discussion

In this study, we investigate neural activity during passive listening to auditory tone sequences by manipulating respective entropy levels and thereby the predictability of an upcoming sound. We used MVPA applied to MEG data to show that, next to more abstract features such as the entropy level, neural responses contain sufficient information to decode the carrier frequency of tones. Our main result reveals that single-trial brain responses to unexpected omissions in a predictable (low entropy) context can be used to decode carrier frequency when a sound is presented. This study provides strong support that top-down prediction related processes are sharply tuned to contain tonotopically specific information. While the finding of sharp tuning of neural activity is not surprising, given in particular invasive recordings from the animal auditory cortex (e.g. during vocalizations, see ((Eliades and Wang, 2003, 2008); or shift of tuning curves following explicit manipulations of attention to specific tone frequencies, see (Fritz et al., 2007a, 2007b)), our work is a critical extension of previous human studies for which a tonotopically tuned effect of predictions has not been shown so far. Critically, given that omission responses have been considered as pure prediction signals (Bendixen et al., 2009; Heilbron and Chait, 2017), our work illustrates that sharp tuning via predictions does not require bottom-up thalamocortical drive.

Sound-evoked activity decreases with increasing regularity, omission-evoked activity increases

As a general test of our data quality, we first focused on evoked responses to pursue some previously reported findings (Alho et al., 1998; Andreou et al., 2015; Raij et al., 1997; Wacongne et al., 2011). Both omissions and sounds elicit the largest evoked responses in the right primary auditory cortex independently from tone-frequency and entropy level. Interestingly, in contrast to sounds, expected but omitted sounds appear to elicit marked evoked activity in the visual cortex. We speculate that unexpected omissions constitute salient events that require reorienting (Corbetta et al., 2008), thereby phase resetting visual activity (for general evidence for audiovisual phase resetting in humans see e.g. (Mercier et al., 2013; Romei et al., 2012)). This potentially interesting question is, however, outside the scope of this manuscript and would require further follow-up studies. Most importantly, sounds and omissions show differential evoked response patterns depending on the contextual entropy level, in particular during later periods of the evoked response (>200 ms). For sound-evoked brain responses amplitude increases with entropy, whereas for omission-evoked brain responses amplitude decrease with entropy. While the omission evoked effect was maximal in left A1, the sound evoked effect was more widespread involving also the striatum and precuneus. The fact that the latter effect goes beyond auditory regions is not surprising given that activity in these regions has been reported to be modulated based on manipulations of regularity in previous studies (Furl et al., 2011; Geiser et al., 2012). For example Rauschecker (Rauschecker, 2011) ascribes the basal ganglia along with other dorsal stream auditory regions a role in matching sounds with expectations formed by previous presentations. Overall, our analysis of evoked responses are fully consistent with previous works (Barascud et al., 2016) and notions of precision based

predictive coding (Heilbron and Chait, 2017), which suggests that neural responses decrease to expected events whereas they increase to unexpected events.

Single-trial MEG activity contains low- and high-level auditory information

To pursue our main research question, we relied on MVPA applied to MEG data (Grootswagers et al., 2017; King and Dehaene, 2014). In particular, prior to addressing whether neural activity during omissions contains carrier-frequency specific information, it was important to illustrate the decoding analysis performance when a sound was actually presented. A priori, this is not a trivial undertaking given that the small spatial extent of the auditory cortex (Saenz and Langers, 2014) likely produces highly correlated topographical patterns for different pure tones and the fact that mapping tonotopic organization using noninvasive electrophysiological tools has had mixed success (for critical overview see e.g. (Lütkenhöner et al., 2003)). Considering this challenging background it is remarkable that *all* participants showed a stable pattern with marked post-stimulus onset decoding increases after ~90 ms. This pattern was observed for all entropy levels and encompassed an on-diagonal increase fading out after ~300 ms and a temporally stable off-diagonal increase indicating a generalizable pattern emerging after ~100 ms and remaining elevated for ~100-200 ms. While this analysis included all sensors and was therefore spatially agnostic, it hints at a rich dynamic that goes beyond a transient activation of a circumscribed brain region (e.g. A1; see also below). Overall, this finding underlines that noninvasive electrophysiological methods such as MEG can be used to decode low-level auditory features such as the carrier frequency of tones. This corroborates and extends findings from the visual modality for which successful decoding of low-level stimulus features such as contrast edge orientation have been demonstrated previously (Cichy et al., 2015).

Going beyond this low-level information, we also addressed whether a representation of a more abstract feature such as the sequence's entropy level could also be decoded from the noninvasive data. Functionally, extracting regularities requires an integration over a longer time period and previous MEG works focussing on evoked responses have identified in particular slow (DC) shifts to reflect transitions from random to regular sound sequences (Barascud et al., 2016). This fits with our result showing that the entropy level of a sound sequence can be decoded above chance at virtually any time point, implying an ongoing (slow) process tracking regularities that is transiently increased following the presentation of a sound. Taken together, the successful decoding of low- and high-level auditory information underlines the significant potential of applying MVPA tools to noninvasive electrophysiological data to address research questions in auditory cognitive neuroscience that would be difficult to pursue using conventional approaches (Crosse et al., 2015; Leonard et al., 2016).

Predictions can be formed in spectrally sharply-tuned manner

Using an MVPA approach with time generalization allowed us to assess whether beyond the level of differential evoked responses, carrier frequency related neural activity during sound

or omission is systematically modulated by the entropy level. In both cases, a clear post-stimulus onset activity pattern was obtained that exhibited a linear relationship to entropy level across participants in the sense that increasing regularity (i.e. lower entropy) went along with improved decoding accuracy. In particular, when training *and* testing on sounds, a pattern emerged after ~150 ms that generalized until ~300 ms and putatively involved left-dominant auditory and non-auditory regions. Note that this effect for sound-to-sound decoding was obtained while overall evoked response strength decreased with increasing regularity. This effect is reminiscent of findings in the visual modality, suggesting a sharpening of the neural response profile by expectations, that is, a reduction of neural responses in the visual cortex, while at the same time representational information is enhanced (Kok et al., 2012). Via our time-generalized omission-to-sound MVPA, we can assert that predictions can sharply tune relevant neural ensembles in a purely top-down manner, without any confounding influence of a sound. In this case an increasing regularity of the sound sequence went along with better decoding for a time period ~100-250 ms training time generalizing to ~200-300 ms testing time. This finding supports and extends a previous experiment by Sanmiguel et al. (Sanmiguel et al., 2013) showing omission responses are sensitive not only to timing, but also to the precise features of the stimulus. Interestingly, the informative time period during the omission is clearly earlier than the omission-evoked response peak and also the period yielding a relationship to entropy level. Also with regards to the latter omission-evoked effect, which was mainly pronounced in left A1, the time-generalized MVPA effect showed a right hemispheric dominance. Altogether, the results underline the fact that our decoding approach uncovers patterns in the data that are not immediately available from looking at the evoked responses. It is worth noting that conforming to the general pattern in this study, the omission-to-sound decoding effect is not confined strictly to auditory regions, but also encompasses (pre-)motor and frontal regions as well as subcortical regions such as the thalamus and hippocampus. This conforms to previous studies in the visual modality implying an involvement of medial temporal and prefrontal regions in the generation of predictions based on the statistical regularity of sensory input (Bar et al., 2006; Schapiro et al., 2012; Summerfield et al., 2006). The differential lateralization patterns for the sound-to-sound and omission-to-sound linear trend effects may be surprising at first sight, especially for the auditory cortex, if one assumes them to reflect pure prediction responses as suggested previously by some authors (Wacongne et al., 2012). However, differential patterns could make sense considering that the absence of an expected stimulation will create a greater amount of surprise than when a sound is presented as expected. This difference could in principle involve a non-overlapping set of brain regions. In any case, this aspect does not change the fact that the late omission-to-sound time-generalized MVPA effect is driven exclusively by top-down processes, illustrating for the first time in humans that prediction-related processes in the auditory system can be tonotopically tuned.

While later latencies could in principle contain a complex mix of prediction and surprise related processes, the act of predicting usually contains a notion of pre-activating relevant neural ensembles, a pattern that has been previously illustrated in the visual modality (e.g. (Ekman et al., 2017; Kok et al., 2017)). For the omission response, this was put forward by Bendixen et al. (Bendixen et al., 2009) even though the reported evoked response effects cannot be directly seen as signatures of preactivation. We found neural patterns ~90 ms

following sound onset to generalize to pre-sound periods that could not be explained by a simple linear carry-over effect from the previous sound. It is thus most parsimonious to assume that next to the activity related to the carrier frequency of the previous tone, pre-sound periods contain relevant information about the carrier frequency of the upcoming tone. However, future studies will need to study this in greater detail since the current design cannot, for example, completely exclude a reactivation of patterns of previous neural activity by a new sound, even though this is not the most parsimonious assumption for the described regression residual effects. Next to this caveat, overall decoding accuracy was in absolute terms not high especially for the critical analysis (i.e. Omission-to-sound decoding). However, it should be noted that we refrained from a widespread practice of subaveraging trials (Cichy et al., 2015; Grootswagers et al., 2017), which boosts classification accuracies significantly. When compared to cognitive neuroscientific M/EEG studies that perform decoding on the genuine single trials and a focus on group level effects (rather than feature-optimizing on individual level as in BCI applications), the strength of our effects are comparable (e.g. (Kaiser et al., 2016; Marti and Dehaene, 2017)).

Acknowledgments

We thank Dr. Anne Hauswald, Dr. Anne Weise and Marta Scislo for helpful comments on earlier versions of the ms, and Miss Hayley Prins for proof reading it. We thank Mr. David Opferkuch and Mr. Manfred Seifter for the help with the measurements.

Author Contributions

Conceptualization, G.D and N.W.; Methodology, G.D., G.S. and N.W.; Investigation and Analysis, G.D., G.S. and N.W.; Writing – Review & Editing, G.D., G.S. and N.W.; Visualization, G.S. and N.W.; Project Administration and Supervision, N.W.

Declaration of Interests

The authors declare no competing interests.

Figure titles and legends

Inline in the main text for better review process

STAR Methods

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Fieldtrip	Oostenveld et al., 2011	http://www.fieldtriptoolbox.org/
MATLAB	MathWorks	https://www.mathworks.com/products/matlab.html
CoSMoMVPA	Oosterhof et al., 2016	http://www.cosmomvpa.org/
Psychophysics Toolbox	Brainard 1997	http://psychtoolbox.org/

Contact for Reagent and Resource Sharing

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Gianpaolo Demarchi (gianpaolo.demarchi@sbg.ac.at).

Experimental Model and Subject Details

Participants

A total of 34 volunteers (16 females) took part in the experiment, giving written informed consent. At the time of the experiment, the average age was 26.6 +- 5.6 SD years. All participants reported no previous neurological or psychiatric disorder, and reported normal or corrected-to-normal vision. The experimental protocol was approved by the ethics committee of the University of Salzburg and has been carried out in accordance with the Declaration of Helsinki.

Method Details

Stimuli and experimental procedure

Before entering the Magnetoencephalography (MEG) cabin, five head position indicator (HPI) coils were applied on the scalp. Anatomical landmarks (nasion and left/right

pre-auricular points), the HPI locations, and around 300 headshape points were sampled using a Polhemus FASTTRAK digitizer. After a 5 min resting state session (not reported in this study), the actual experimental paradigm started. The subjects watched a movie (*Cirque du Soleil: Worlds Away)* while passively listening to tone sequences. Auditory stimuli were presented binaurally using MEG-compatible tubal in-ear headphones (SOUNDPixx, VPixx technologies, Canada). This particular movie was chosen for the absence of speech and dialogue, and the soundtrack was substituted with the sound stimulation sequences. These sequences were composed of four different pure (sinusoidal) tones, ranging from 200 to 2000 Hz, logarithmically spaced (that is: 200 Hz, 431 Hz, 928 Hz, 2000 Hz) each lasting 100 ms (5 ms linear fade in / out). Tines were presented at a rate of 3 Hz. Overall the participants were exposed to four blocks, each containing 4000 stimuli, with every block lasting about 22 mins. Each block was balanced with respect to the number of presentations per tone frequency. Within the block, 10% of the stimuli were omitted, thus yielding 400 omission trials (100 per omitted sound frequency). While within each block, the overall amount of trials per sound frequency was set to be equal, blocks differed in the order of the tones, which were parametrically modulated in their entropy level using different transition matrices (Nastase et al., 2014). In more detail, the random condition (RND; see **Figure 1**) was characterized by equal transition probability from one sound to another, thereby preventing any possibility of accurately predicting an upcoming stimulus (high entropy). In the ordered condition (OR), presentation of one sound was followed with high (75%) probability by another sound (low entropy). Furthermore, two intermediate conditions were included (MM and MP). The probability on the diagonal was set to be equiprobable (25%) across all entropy conditions, thereby controlling for the influence of self-repetitions. The experiment was programmed in MATLAB 9.1 (The MathWorks, Natick, Massachusetts, U.S.A) using the open source Psychophysics Toolbox (Brainard, 1997).

MEG data acquisition and preprocessing

The magnetic signal was recorded at 1000 Hz (hardware filters: 0.1 - 330 Hz) in a standard passive magnetically shielded room (AK3b, Vacuumschmelze, Germany) using a whole head MEG (Elekta Neuromag Triux, Elekta Oy, Finland). Signals were sampled with 102 magnetometers and 204 orthogonally placed planar gradiometers at 102 different positions. We use a signal space separation algorithm implemented in the Maxfilter program (version 2.2.15) provided by the MEG manufacturer to remove external noise from the MEG signal (mainly 16.6Hz, and 50Hz plus harmonics) and realign data to a common standard head position (*-trans default* Maxfilter parameter) across different blocks based on the measured head position at the beginning of each block (Cichy and Pantazis, 2017).

Data analysis was done using the Fieldtrip toolbox (git version 20170919; (Oostenveld et al., 2011)) and in-house built scripts. First, a high-pass filter at 0.1 Hz (6th order zero-phase Butterworth filter) was applied to the continuous data. Then the data were segmented from 600 ms before to 600 ms after target stimulation onset and down-sampled to 256 Hz for the ERF analysis, and to 100 Hz for the decoding part. Trials containing physiological or acquisition artifacts were rejected. A semi-automatic artifact detection routine identified

statistical outliers of trials in the datasets using a set of summary statistics (variance, maximum absolute amplitude, maximum z-value). These trials were removed from each dataset. Across subjects, an average of 721 ± 266 SD (4.5 ± 1.7 SD %) of trials were rejected. In all further analyses for each subject, the number of trials for the different carrier frequencies was balanced to prevent any bias across conditions (Gross et al., 2013). Finally, the epoched data was 30 Hz lowpass-filtered (6th order zero-phase Butterworth filter) prior to further analysis.

Source level analysis

Preprocessed data was projected to source-level using an LCMV beamformer analysis (Van Veen et al., 1997). For each participant, realistically shaped, single-shell headmodels (Nolte, 2003) were computed by co-registering the participants' headshapes either with their structural MRI (15 participants) or – when no individual MRI was available (19 participants) – with a standard brain from the Montreal Neurological Institute (MNI, Montreal, Canada), warped to the individual headshape. A grid with 1 cm resolution based on an MNI template brain was morphed to fit the brain volume of each participant. A common spatial filter (for each grid point and each participant) was computed using the leadfields and the common covariance matrix, taking into account the data from all trials (i.e. including sound and omission trials from all conditions). The covariance window for the beamformer filter calculation was based on 200 ms pre-stimulus to 500 ms post-stimulus. Using this common filter, the sensor level single-trial time-series were projected onto the 3D grid. For the evoked response, the resulting sound and omission trials were averaged relative to the stimulus onset and the absolute value was calculated. This yields for each condition a sound- or omission-related amplitude time series. Baseline normalization was performed only for visualization purposes (subtraction of 50-ms pre-stimulus activity).

Multivariate Pattern Analysis (MVPA)

We used multivariate pattern analysis as implemented in CoSMoMVPA (git version 20170505; (Oosterhof et al., 2016)). MVPA decoding was first performed using a time-generalized decoding analysis that included all magnetometers. Specific time slices from the time-generalization matrix were followed up by a spatial-searchlight decoding at source level (see below). We performed decoding analysis based on single trial sensor-level data and single trial normalized (z-scored) source data.

Overall, three decoding approaches were taken:

- *Entropy-level decoding*: In a first step, we kept only trials with sound presentation (removing omission trials) to investigate brain activity modulated by different experimental contexts. For this purpose, we defined four decoding targets (classes) based on block type (4 contexts: RND, MM, MP, OR).
- *Sound-to-sound decoding*: To test whether we could classify carrier frequency in general, we defined four targets (classes) for the decoding related to the carrier frequency of the sound presented on each trial (4 carrier frequencies).
- *Omission-to-sound decoding*: To test whether omission periods contain carrier frequency specific neural activity, omission trials were labeled according to the carrier frequency of the sound which would have been presented. These trials were used to

train the classifier, which was subsequently applied to a test set of trials during which sounds were presented.

Using a Linear Discriminant Analysis (LDA) classifier, we performed a decoding analysis at each time point around stimulus / omission onset. A two-fold cross-validation scheme was applied for entropy-level and sound-to-sound decoding, using two randomly assigned sets of single trials. For the omission-to-sound decoding analysis, the training set was restricted to omission trials and the testing set contained only sound trials. Trials were balanced in the training and testing sets by using a random subset of trials in which the number of trials was equalized between the four conditions (i.e. 4 target classes: 4 entropy levels or 4 carrier frequencies depending on the decoding analysis). In all cases, training and testing partitions always contained different sets of data.

Classification accuracy for each subject was averaged at the group level and reported to depict the classifier's ability to decode over time (i.e. time-generalization analysis at sensor level) and over spatial dimension (i.e. searchlight analysis at source level). The time generalization method was used to study the ability of each LDA classifier across different time points in the training set to generalize to every time point in the testing set (King and Dehaene, 2014). For the sound-to-sound and omission-to-sound decoding, time generalization was calculated for each entropy level separately, resulting in four generalization matrices, one for each entropy level. This was necessary to assess whether the contextual sound sequence influences classification accuracy on a systematic level. Significant clusters of time points were followed up by a searchlight analysis across brain sources. In this analysis we used local neighborhood features in source space (source radius of 1.5 cm). All significant searchlight accuracy results were averaged over time cluster and reported on brain maps.

Quantification and Statistical analysis

For the evoked responses, we tested the dependence on entropy level using a regression test (*depsamplesregT* in Fieldtrip). Results for sounds and omissions were sorted from random to ordered respectively. Testing sound- and omission-evoked responses separately on a whole brain level first, we defined an early (100-200 ms) and late time window (200-300 ms) based on previous studies in this domain (for an overview see (Heilbron and Chait, 2017; Schröger et al., 2015)). In order to account for multiple comparisons, we used a nonparametric cluster permutation test (Maris and Oostenveld, 2007) as implemented in Fieldtrip using 1000 permutations and a $p < .025$ to threshold the clusters. Neighboring grid points were clustered (minimum number of grid points in a cluster when their distance was below 1.5 cm). Given previous works (Alink et al., 2010; Todorovic and de Lange, 2012) and also theoretical reasoning (Summerfield and de Lange, 2014), we hypothesized decreasing evoked responses to sounds the more ordered the sound sequence became. On the other hand, for omissions, we expected evoked responses to increase the more ordered the

sequences became, since within these sequences expectations and violations thereof should be stronger (Sanmiguel et al., 2013) This latter prediction was not evident at a whole brain cluster corrected level. In order to target this differential prediction for sound and omission evoked responses in a more direct manner, we implemented a normalized contrast and focused on the left and right auditory cortex (as given by the grand average; see **Figure 2A**). In this procedure, we first normalized each condition (i.e. sound / omission x entropy level) by the evoked response of the random sequence (e.g. OR*norm* = [OR – RND] / [OR + RND]). For the regression analysis, we entered the difference of the normalized contrasts between omission and sound (e.g. OR*diff* = [OR*norm* (omission) - OR*norm* (sound)]). According to our hypothesis, the differential relationship to entropy level for sound and omission evoked responses should be reflected in a monotonically increasing difference.

The multivariate analysis results were tested at the group level by comparing the resulting individual accuracy maps against chance level (25% with 4 classes) using a non-parametric approach implemented in CoSMoMVPA (Oosterhof et al., 2016) adopting 10,000 permutations to generate a null distribution. P-values were set at $p < 0.005$ for cluster level correction to control for multiple comparisons using a threshold-free method for clustering (Smith and Nichols, 2009), which has been used and validated for MEG/EEG data (Pernet et al., 2015; Turella et al., 2016). The time generalization results and searchlight brain maps at the group level were thresholded using a mask with corrected z-score > 2.58 (or p corrected < 0.005). We also tested the dependence of classification results on entropy level using a regression test (*depsamplesregT* in Fieldtrip) following analogous statistical method as evoked response analysis. Only the significant time-by-time points identified on sensor level time-generalization where used to test source level dependence of searchlight decoding results on entropy level using a similar regression test.

Data and software availability

Further information and requests for resources or data should be directed to and will be fulfilled by the corresponding author..

References

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jääskeläinen, I.P., Pekkonen, E., and Ilmoniemi, R.J. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. Psychophysiology *35*, 211–224.

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., and Muckli, L. (2010). Stimulus

predictability reduces responses in primary visual cortex. J. Neurosci. *30*, 2960–2966.

Andreou, L.-V., Griffiths, T.D., and Chait, M. (2015). Sensitivity to the temporal structure of rapid sound sequences - An MEG study. Neuroimage *110*, 194–204.

Auksztulewicz, R., Barascud, N., Cooray, G., Nobre, A.C., Chait, M., and Friston, K. (2017). The Cumulative Effects of Predictability on Synaptic Gain in the Auditory Processing Stream. J. Neurosci. *37*, 6751–6760.

Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Schmidt, A.M., Dale, A.M., Hämäläinen, M.S., Marinkovic, K., Schacter, D.L., et al. (2006). Top-down facilitation of visual recognition. Proc. Natl. Acad. Sci. U. S. A. *103*, 449–454.

Barascud, N., Pearce, M.T., Griffiths, T.D., Friston, K.J., and Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. Proc. Natl. Acad. Sci. U. S. A. *113*, E616–E625.

Bendixen, A., Schröger, E., and Winkler, I. (2009). I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. J. Neurosci. *29*, 8447–8451.

Bendixen, A., Scharinger, M., Strauß, A., and Obleser, J. (2014). Prediction in the service of comprehension: modulated early brain responses to omitted speech segments. Cortex *53*, 9–26.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. *10*, 433–436.

Chennu, S., Noreika, V., Gueorguiev, D., Shtyrov, Y., Bekinschtein, T.A., and Henson, R. (2016). Silent Expectations: Dynamic Causal Modeling of Cortical Prediction and Attention to Sounds That Weren't. J. Neurosci. *36*, 8305–8316.

Cichy, R.M., and Pantazis, D. (2017). Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. Neuroimage *158*, 441–454.

Cichy, R.M., Ramirez, F.M., and Pantazis, D. (2015). Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? Neuroimage *121*, 193–204.

Corbetta, M., Patel, G., and Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. Neuron *58*, 306–324.

Crosse, M.J., Butler, J.S., and Lalor, E.C. (2015). Congruent Visual Speech Enhances Cortical Entrainment to Continuous Auditory Speech in Noise-Free Conditions. J. Neurosci. *35*, 14195–14204.

Ekman, M., Kok, P., and de Lange, F.P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. Nat. Commun. *8*, 15276.

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

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

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the

primate cerebral cortex. Cereb. Cortex *1*, 1–47.

Friston, K. (2005). A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. *360*, 815–836.

Fritz, J.B., Elhilali, M., David, S.V., and Shamma, S.A. (2007a). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? Hear. Res. *229*, 186–203.

Fritz, J.B., Elhilali, M., and Shamma, S.A. (2007b). Adaptive changes in cortical receptive fields induced by attention to complex sounds. J. Neurophysiol. *98*, 2337–2346.

Furl, N., Kumar, S., Alter, K., Durrant, S., Shawe-Taylor, J., and Griffiths, T.D. (2011). Neural prediction of higher-order auditory sequence statistics. Neuroimage *54*, 2267–2277.

Geiser, E., Notter, M., and Gabrieli, J.D.E. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. J. Neurosci. *32*, 6177–6182.

Grootswagers, T., Wardle, S.G., and Carlson, T.A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. J. Cogn. Neurosci. *29*, 677–697.

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., et al. (2013). Good practice for conducting and reporting MEG research. Neuroimage *65*, 349–363.

Heilbron, M., and Chait, M. (2017). Great expectations: Is there evidence for predictive coding in auditory cortex? Neuroscience.

Jaramillo, S., and Zador, A.M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. Nat. Neurosci. *14*, 246–251.

Kaiser, D., Oosterhof, N.N., and Peelen, M.V. (2016). The Neural Dynamics of Attentional Selection in Natural Scenes. J. Neurosci. *36*, 10522–10528.

King, J.-R., and Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn. Sci. *18*, 203–210.

Kok, P., Jehee, J.F.M., and de Lange, F.P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. Neuron *75*, 265–270.

Kok, P., Mostert, P., and de Lange, F.P. (2017). Prior expectations induce prestimulus sensory templates. Proc. Natl. Acad. Sci. U. S. A. *114*, 10473–10478.

Kraemer, D.J.M., Macrae, C.N., Green, A.E., and Kelley, W.M. (2005). Musical imagery: sound of silence activates auditory cortex. Nature *434*, 158.

Leonard, M.K., Baud, M.O., Sjerps, M.J., and Chang, E.F. (2016). Perceptual restoration of masked speech in human cortex. Nat. Commun. *7*, 13619.

Lütkenhöner, B., Krumbholz, K., and Seither-Preisler, A. (2003). Studies of tonotopy based on wave N100 of the auditory evoked field are problematic. Neuroimage *19*, 935–949.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and

MEG-data. J. Neurosci. Methods *164*, 177–190.

Marti, S., and Dehaene, S. (2017). Discrete and continuous mechanisms of temporal selection in rapid visual streams. Nat. Commun. *8*, 1955.

Mercier, M.R., Foxe, J.J., Fiebelkorn, I.C., Butler, J.S., Schwartz, T.H., and Molholm, S. (2013). Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. Neuroimage *79*, 19–29.

Müller, N., Keil, J., Obleser, J., Schulz, H., Grunwald, T., Bernays, R.-L., Huppertz, H.-J., and Weisz, N. (2013). You can't stop the music: reduced auditory alpha power and coupling between auditory and memory regions facilitate the illusory perception of music during noise. Neuroimage *79*, 383–393.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin. Neurophysiol. *118*, 2544–2590.

Nastase, S., Iacovella, V., and Hasson, U. (2014). Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. Hum. Brain Mapp. *35*, 1111–1128.

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. Phys. Med. Biol. *48*, 3637–3652.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 156869.

Oosterhof, N.N., Connolly, A.C., and Haxby, J.V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. Front. Neuroinform. *10*, 27.

Pernet, C.R., Latinus, M., Nichols, T.E., and Rousselet, G.A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. J. Neurosci. Methods *250*, 85–93.

Plakke, B., and Romanski, L.M. (2014). Auditory connections and functions of prefrontal cortex. Front. Neurosci. *8*, 199.

Raij, T., McEvoy, L., Mäkelä, J.P., and Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. Brain Res. *745*, 134–143.

Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. *2*, 79–87.

Rauschecker, J.P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. Hear. Res. *271*, 16–25.

Rauschecker, J.P., and Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. *97*, 11800–11806.

Recasens, M., and Uhlhaas, P.J. (2017). Test-retest reliability of the magnetic mismatch

negativity response to sound duration and omission deviants. Neuroimage *157*, 184–195.

Romei, V., Gross, J., and Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. Curr. Biol. *22*, 807–813.

Rubin, J., Ulanovsky, N., Nelken, I., and Tishby, N. (2016). The Representation of Prediction Error in Auditory Cortex. PLoS Comput. Biol. *12*, e1005058.

Saenz, M., and Langers, D.R.M. (2014). Tonotopic mapping of human auditory cortex. Hear. Res. *307*, 42–52.

Sanmiguel, I., Saupe, K., and Schröger, E. (2013). I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted “what” but not “when.” Front. Hum. Neurosci. *7*, 407.

Schapiro, A.C., Kustner, L.V., and Turk-Browne, N.B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. Curr. Biol. *22*, 1622–1627.

Schröger, E., Marzecová, A., and SanMiguel, I. (2015). Attention and prediction in human audition: a lesson from cognitive psychophysiology. Eur. J. Neurosci. *41*, 641–664.

Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage *44*, 83–98.

Spratling, M.W. (2017). A review of predictive coding algorithms. Brain Cogn. *112*, 92–97.

Summerfield, C., and de Lange, F.P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. Nat. Rev. Neurosci. *15*, 745–756.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., and Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. Science *314*, 1311–1314.

Todorovic, A., and de Lange, F.P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. J. Neurosci. *32*, 13389–13395.

Todorovic, A., van Ede, F., Maris, E., and de Lange, F.P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J. Neurosci. *31*, 9118–9123.

Turella, L., Tucciarelli, R., Oosterhof, N.N., Weisz, N., Rumiati, R., and Lingnau, A. (2016). Beta band modulations underlie action representations for movement planning. Neuroimage *136*, 197–207.

Ungerleider, L. (1994). What and Where in the Human Brain. Curr. Opin. Neurobiol. *4*, 157–165.

Van Veen, B.D., van Drongelen, W., Yuchtman, M., and Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. *44*, 867–880.

Voisin, J., Bidet-Caulet, A., Bertrand, O., and Fonlupt, P. (2006). Listening in silence activates auditory areas: a functional magnetic resonance imaging study. J. Neurosci. *26*,

273–278.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., and Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc. Natl. Acad. Sci. U. S. A. *108*, 20754–20759.

Wacongne, C., Changeux, J.-P., and Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. J. Neurosci. *32*, 3665–3678.

Whitford, T.J., Jack, B.N., Pearson, D., Griffiths, O., Luque, D., Harris, A.W., Spencer, K.M., and Le Pelley, M.E. (2017). Neurophysiological evidence of efference copies to inner speech. Elife *6*.

Wolpert, D.M., and Miall, R.C. (1996). Forward Models for Physiological Motor Control. Neural Netw. *9*, 1265–1279.

Wolpert, D.M., Ghahramani, Z., and Jordan, M.I. (1995). An internal model for sensorimotor integration. Science *269*, 1880–1882.