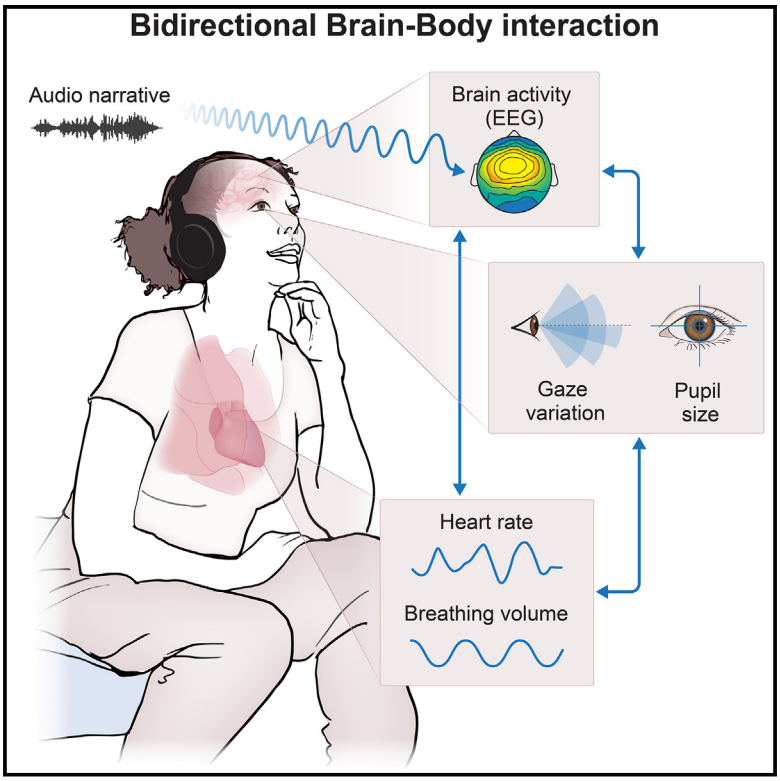
Cell Reports

Article

Bidirectional brain-body interactions during natural story listening

Graphical abstract



Authors

Jens Madsen, Lucas C. Parra

Correspondence

jmadsen@ccny.cuny.edu

In brief

When we listen to a story, our body is integrally involved in the experience. Madsen and Parra provide evidence for a bidirectional link between body and brain by analyzing electroencephalography, pupil size, heart rate, and eye movements while subjects listen to narratives and during controlled interventions that modulate autonomic signals.

Highlights

- Listening to narratives modulates eye movements
- Heart rate fluctuations precede pupil size fluctuations and anterior-central neural activity
- Breathing modulates pupil size, suggesting causal effect on central arousal
- Rhythmic saccades can entrain heart beats

Madsen & Parra, 2024, Cell Reports *43*, 114081
April 23, 2024
https://doi.org/10.1016/j.celrep.2024.114081

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Bidirectional brain-body interactions during natural story listening

Jens Madsen[1,2,*] and Lucas C. Parra[1]
[1]Department of Biomedical Engineering, City College of New York, 85 St. Nicholas Terrace, New York, NY 10031, USA
[2]Lead contact
*Correspondence: jmadsen@ccny.cuny.edu
https://doi.org/10.1016/j.celrep.2024.114081

SUMMARY

Narratives can synchronize neural and physiological signals between individuals, but the relationship between these signals, and the underlying mechanism, is unclear. We hypothesized a top-down effect of cognition on arousal and predicted that auditory narratives will drive not only brain signals but also peripheral physiological signals. We find that auditory narratives entrained gaze variation, saccade initiation, pupil size, and heart rate. This is consistent with a top-down effect of cognition on autonomic function. We also hypothesized a bottom-up effect, whereby autonomic physiology affects arousal. Controlled breathing affected pupil size, and heart rate was entrained by controlled saccades. Additionally, fluctuations in heart rate preceded fluctuations of pupil size and brain signals. Gaze variation, pupil size, and heart rate were all associated with anterior-central brain signals. Together, these results suggest bidirectional causal effects between peripheral autonomic function and central brain circuits involved in the control of arousal.

INTRODUCTION

As we engage with the world, we move our eyes, scanning our surroundings. Fluctuations in lighting cause the pupil to contract or dilate,[1] and even in the absence of luminance changes, our pupil fluctuates.[2] During task demand, our autonomic system regulates dilation and constrictions of the pupil via sympathetic and parasympathetic activity.[3] A major hub of the autonomic system that regulates arousal is the locus coeruleus (LC).[4] Its activity is tightly linked to pupil size,[5] and thus pupil size is often considered a marker of arousal.[6] These task-evoked pupil responses have been shown to entrain to task timing,[7] cognitive control,[8] and cognitive resource allocation in mental arithmetic tasks.[9] Interestingly, a number of cognitive factors modulate pupil size, such as affective processing,[10–12] cognitive effort,[13] cognitive control,[14] task engagement,[15] conscious processing,[16] attention,[17,18] and more. This suggests a direct effect of cognition on arousal. Phasic fluctuations of LC neurons have been associated with task-related decision processes, whereas tonic activity has been associated with disengagement from the current task and a search for alternative behaviors.[19]

The heart is also controlled by the autonomic nervous system, with sympathetic response increasing heart rate and parasympathetic activity decreasing it.[20] When we are resting, heart rate fluctuates with respiration, which is known as respiratory sinus arrhythmia.[21] Stressful mental effort can increase heart rate.[22] Even subtle arousal due to emotional[23] or surprising[24] visual stimuli can accelerate individual heart beats. At a slower timescale, heart rate is positively correlated with cortical activity in fMRI, predominantly in midline brain areas.[25] The overall strength of fluctuations—measured as heart rate variability[26]—is affected by various cognitive factors.[27–32] We recently demonstrated that audiovisual narratives can reliably increase or decrease heart rate at different moments of a story and thus synchronize heart rate across people.[33] This suggests that the cognitive processing of the narrative has an immediate effect on heart rate.

Eye movements have also been linked to arousal in a recent study,[15] which argues that a common subcortical signal controls pupil dilation as well as saccade initiation. We have recently found that engaging videos reliably modulate saccade rate (the number of saccades per second) but only if viewers are paying attention.[34] This suggests that cognitive processing of the visual stimulus modulates saccade rate. What is not yet known is whether a purely auditory narrative will also affect pupil size and eye movements, as we have observed for heart rate.[33] If heart rate fluctuated as a result of fluctuating arousal, we might expect that the pupil and saccades will be similarly driven by an auditory-only narrative. Here, we hypothesized that the cognitive processing of an auditory narrative will drive not only heart rate but also pupil size and eye movements.

To assess this, we measured inter-subject correlation (ISC) of these peripheral signals, as well as scalp potentials. ISC is an established metric to capture processing associated with dynamic natural stimuli such as films[35] or auditory narratives.[36,37] Scalp potentials are known to entrain to low-level features of continuous speech[38,39] as well as its semantics[40] (i.e., "speech tracking"[41]). To disambiguate cognitive processing of the narrative from purely low-level auditory processing, we will use attentive and distracted listening conditions[42] and, additionally, analyze responses to words and their information rate. With these experiments, we aim to test whether there is a top-down

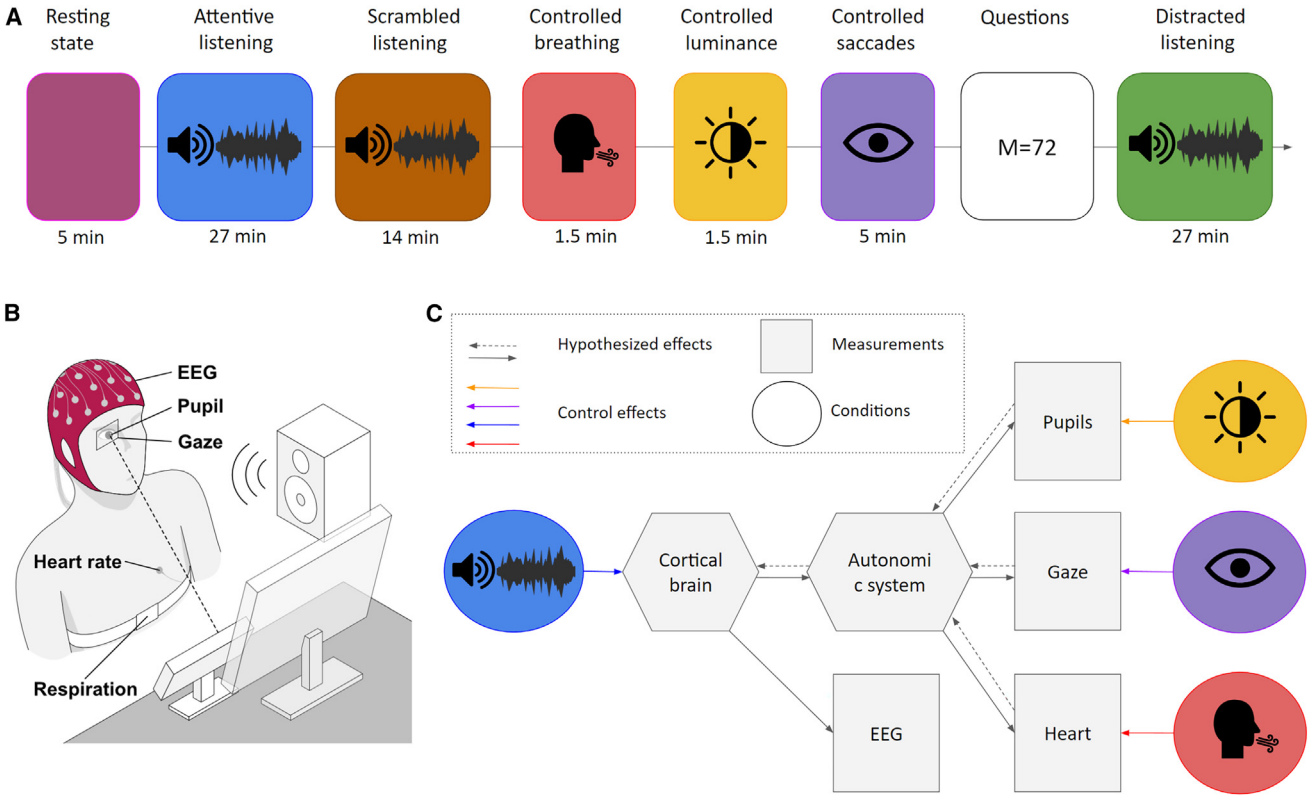


Figure 1. Synchronization of physiological and neural signals while listening to auditory narratives

(A) Experimental paradigm: an initial 5 min of resting is followed by attentive listening to audio narratives (10 stories of 27 min duration) (for experiment 2, participants listened to 5 scrambled versions of the stories, 14 min duration), controlled breathing for 90 s, controlled luminance for 60 s (screen fluctuating between black and white at f = 0.1 Hz), and controlled saccades (dots appear sequentially at 9 points on a grid with varying durations). Then, questions are asked about the stories (4–10 questions per story), and lastly, the subjects were distracted while listening to the same stories (now counting backwards in decrements of 7 silently in their heads).

(B) Signal modalities collected: scalp potentials of electroencephalography (EEG), eye tracking (horizontal, vertical), pupil size, electrocardiography (ECG), and respiration.

(C) Diagram of measurements, controlled effects, and hypothesized effects. Not shown in this conceptual schematic are the direct (artifactual) effects on the EEG from cardiac signal, eye movement, and possibly breathing movement.

effect of cognitive processing of the narrative on pupil size and eye movements, without confounds of a visual stimulus or motor behavior, while controlling for the known effects of eye movements on pupil size.[43,44]

To quantify bottom-up effects, we will use interventions that are known to modulate these peripheral signals. We exploit the fact that heart rate is modulated by breathing,[45] increasing with inhalation and decreasing with exhalation. Pupil size is similarly affected by luminance.[1] Finally, we elicit controlled saccades by asking participants to follow a dot jumping on the screen. We hypothesize that these interventions will result in a modulation of the other peripheral signals. For instance, fluctuating heart rate (as a result of rhythmic breathing) might lead to rhythmic fluctuations of pupil size. Similarly, rhythmic saccades might affect heart rate. Heart rate and eye movements are controlled by inter-connected brainstem nuclei.[15,18,27] If we do find such interactions, then this suggests that there are common brain nuclei centrally driving these signals. Importantly, a modulation of pupil size would suggest a modulation of cortical excitability, given the ample evidence that pupil size reflects bottom-up cortical activation during quiet wakefulness,[46] via the ascending arousal system.[47,48] To summarize, we will explore a top-down effect of cognitive processing of a narrative on pupil size, eye movements, and heart rate and determine if controlled interventions known to modulate individual peripheral signals also modulate other peripheral signals.

RESULTS

We experimentally manipulate four factors at play (Figure 1A): attention to auditory narratives in an attentive (blue) and distracted condition (green), heart rate through controlled breathing (red), gaze position through controlled saccades or fixation cross (purple), and pupil size through controlled luminance (yellow). We also included a rest condition (pink) to capture spontaneous brain-body interactions. In the distracted listening condition, participants were asked to perform a mental arithmetic task, which was meant to distract them from the narrative. To probe memory, toward the end of the experiment, we asked participants questions to test recall of the content of the auditory narrative.

We know that listening to autobiographical narratives will drive cortical activity similarly across participants, measured as ISC of scalp potentials.[37] The effect is strongest during attentive listening and predictive of the recall performance. This suggests that cognitive processing of the narratives drives cortical activity. We hypothesized that this cognitive processing has a top-down effect on the autonomic system affecting not only heart rate[33] but also eye movements and pupil size (hypothesis H1, solid arrows in Figure 1C). We therefore predicted that these peripheral signals and the electroencephalographic (EEG) scalp potentials will be correlated between and within subjects. For brevity, we will refer to this as the ''effect of cognition'' or ''cognitive drive.'' We will test this with signals recorded during the narrative in both attentive and distracted conditions as well as during rest. The mental arithmetic task in the distracted condition also involves cognitive processes that may drive autonomic signals and induce a correlation within subjects, but we do not expect them to be correlated between subjects. Lastly, we hypothesized a bottom-up effect of autonomic physiology on the brain (hypothesis H2, dashed arrows in Figure 1C). We will test this by correlating peripheral signals during conditions controlling luminance, saccades, and breathing in separate control conditions.

A total of $N = 67$ participants were analyzed on data collected in two experiments (47 females, age 18–36, mean = 22.49, standard deviation = 3.71). The first cohort ($n = 29$) looked at a blank gray screen during the narratives and during rest (Figure 1B). The second cohort ($n = 38$) was asked to instead look at a fixation cross with luminance matching the gray screen. All results reported here were reproduced in these two cohorts except where differences are highlighted.

Saccade initiation and blink onset are affected by cognitive processing of the auditory narrative

Hypothesis H1 predicts that eye movements will be driven by the auditory narrative. In an effort to emulate natural listening conditions, for the first cohort of subjects, we did not require anything from the subjects other than to look at a blank gray screen during the listening task.

If the cognitive processing of the narrative elicits saccades, then we would expect them to be initiated at specific moments of the narrative. To test this, we measured how similar saccade initiation was across listeners using ISC. We resolved this in frequency to determine at what timescale of the narrative this occurs. We find that saccade initiation did synchronize between subjects during the attentive listening (Figure 2A, $n = 29$, k = 101, $p < 0.01$ in 0.02–0.13 and 0.23–0.93 Hz). The timescale of synchronization of 0.1 Hz or less is much slower than the average saccade rate of approximately 2 Hz (attentive: 1.79 ± 0.96 Hz, distracted: 1.91 ± 0.78 Hz). This suggests that specific moments of the narrative tended to elicit saccades in multiple subjects, and importantly, this is modulated by attention. We note that the number and the size of saccades during the narratives were not consistently different between the attentive and distracted conditions (Figures S2A and S2B). We also tested if saccade rate fluctuations entrained to the auditory narratives as we had previously observed for video.[49] However, we do not find that the saccade rate changes in power (Figure S2E) or in synchronization between subjects (Figure S2F). Therefore, the basic statistics of saccades are largely unaffected by the narrative.

Previous reports have linked blinking to pauses in continuous speech.[50] To test if we see this effect here and if it depends on cognitive processing, we compute the ISC of blinking (Figure 2B). We find a similar behavior to saccades, namely, blinking is locked to the narrative at a slow timescale (Figure 2B, $n = 29$, k = 101, $p < 0.01$ in 0.02–0.13 and 0.23–1.00 Hz). This is slower than the typical blink rate (0.75 ± 0.7 Hz; see Figure S3D) and the average word rate of the stories (2.90 ± 0.28 words/s; see Figure S6D), so blinks may occasionally be evoked at specific moments in the narrative and could be at the timescale of sentences (0.3 ± 0.29 sentences/s; see Figure S6H). In summary, both saccade initiation and blink onset are driven by the narrative in the absence of a visual stimulus.

Gaze variation is modulated by attention and correlates with anterior-central neural activity

We found that gaze position fluctuated over time. A power spectrum reveals the amplitude of these fluctuations at different timescales (Figure 2C)—we refer to these as gaze variations in the following. Attentive listening resulted in increased gaze variation around 0.1 Hz ($n = 29$, k = 101, $p < 0.01$ in 0.06–0.38 Hz) (see example video of these slow gaze variations here). This is substantially slower than the speed of saccades or their rate of occurrence (Figure S2D, saccade rate M = 1.8 ± 0.74 saccades/s). These slow variations in gaze position suggest that subjects do not saccade more frequently but simply that gaze position covers a wider field of view when attentively listening to the story.

After these findings on gaze variation, we repeated the experiment on a second cohort of participants (experiment 2, $N = 38$), but now we asked listeners to look at a fixation cross (with luminance equal to the background). Surprisingly, these slow gaze variations were still larger during the attentive versus the distracted listening conditions (Figure S5C, $n = 38$, k = 101, $p < 0.01$ in 0.02–0.18 Hz). Only during rest, where the only task was to maintain fixation, did we have the expected drop in gaze variation (Figures S5A and S5B). In short, when participants are asked to focus on a cognitive task (narrative or counting), they do not maintain fixation as well.

To determine if these gaze variations are driven by the narrative, we measured their correlation across subjects during the narrative (Figure 2D). During attentive listening, we see significant ISC in the range of 0.04–0.25 Hz ($N = 29$, k = 101, $p < 0.01$ in 0.04–0.25 Hz), and this is absent during the distracted conditions, resulting in a significant difference between attentive and distracted conditions (Figure 2D, gray shading). This difference was reproduced when subjects were asked to maintain fixation (Figure S5D, $n = 38$, k = 101, $p < 0.01$ in 0.02–0.13 Hz). To control for low-level sound processing more directly, in experiment 2, we also included a condition with noise-like (''scrambled'') audio immediately after the attentive condition. Subjects were better at maintaining fixation in the scrambled condition (Figure 2E), and ISC was significantly lower as compared to the attentive condition for gaze variation (Figure 2F, $n = 38$, k = 101, $p < 0.01$ in 0.02–0.13 Hz), as well as saccades and blinks

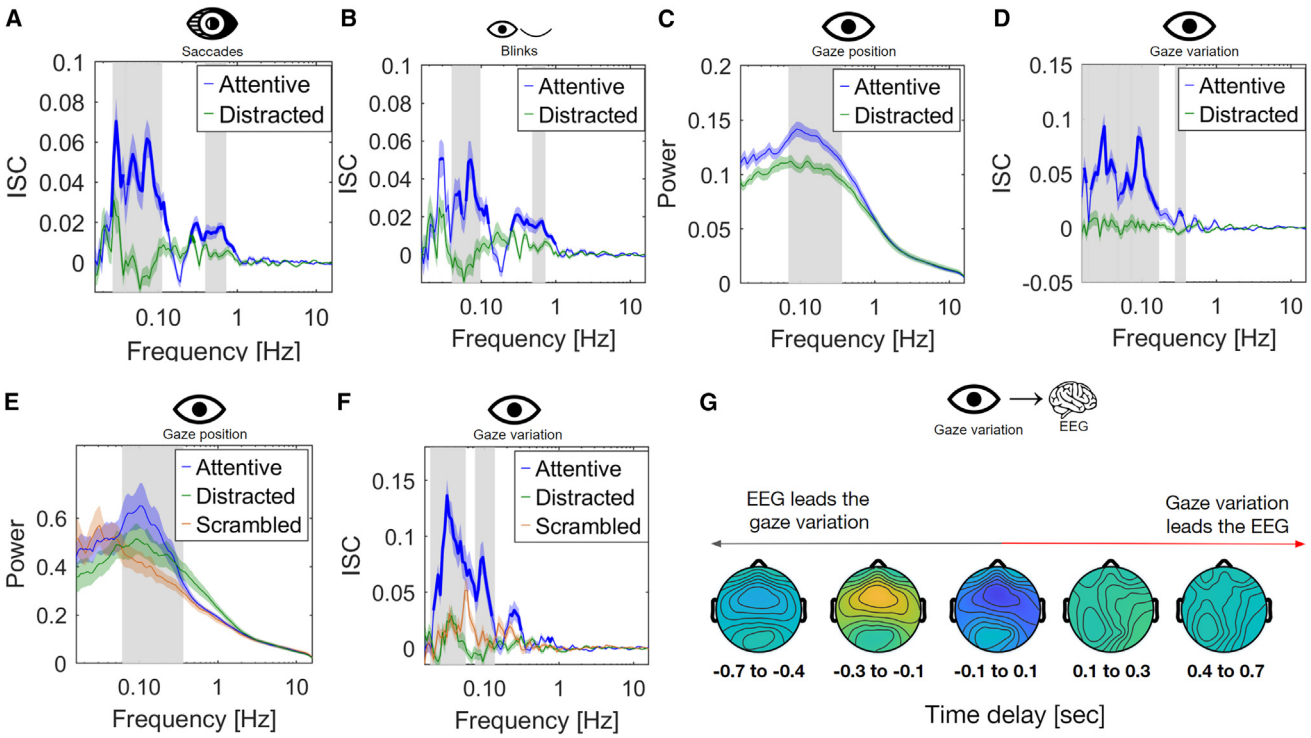


Figure 2. Eye movements during listening to auditory narratives

(A) Frequency-resolved ISC computed by first band-pass filtering the saccade signals (set to 1 during a saccade occurs and 0 otherwise) at different frequencies and then computing ISC averaged over stories and subjects. Colored shading indicates standard error around the mean (SEM) across $n = 29$ subjects. Significant ISC is shown as a bold line. Significant difference in ISC between the attentive and distracted conditions is indicated as gray shading. Significance is established in each band using t-statistics, corrected for multiple comparisons using one-dimensional cluster statistics ($p < 0.01$, $n = 29$, k = 101 frequency bands; see STAR Methods for details).

(B) Frequency-resolved ISC of blinks computed between blink signals (computed as saccades in A). Shading and significance computed as in (A).

(C) Power spectrum of gaze position computed by first band-pass filtering the gaze position signal, computing the standard deviation for each subject, taking the mean over stories and vertical/horizontal direction, and lastly normalizing it by the total power per subject across conditions. Significant difference between the attentive and distracted conditions is established in each band using Wilcoxon rank-sum test, corrected for multiple comparisons using one-dimensional cluster statistics (light gray area indicates $p < 0.01$, cluster corrected).

(D) Frequency-resolved ISC of gaze variation computed as the absolute value of the Hilbert-transformed gaze position after band-pass filtering. Shading and significance computed as in (A).

(E) Same as (C) but for experiment 2 ($n = 38$), including a modulated auditory noise condition (scrambled). Power was computed for the 5 stories shared between the 3 conditions (14 min). Shaded area indicates significant difference between the attentive and scrambled conditions. Significance computed as in (C).

(F) Same as (D) but for experiment 2 ($n = 38$), including a modulated auditory noise condition (scrambled). ISC is computed for the 5 stories shared between the 3 conditions (14 min). Significance computed as in (A).

(G) Temporal response functions (TRF) of scalp potentials using each EEG electrode as output signal and the band-pass-filtered gaze variation signal (0.1–1 Hz) as input signal during the auditory narratives for experiment 1 ($n = 29$). Color indicates scalp potential fluctuations in response to fluctuations of gaze variation (yellow and blue are positive and negative correlations, respectively).

(Figures S2C and S3C). This indicates that the moments of high and low gaze variation coincided across participants but only if they were attending to the narrative.

To measure the relation between these gaze variations and cortical brain activity, we estimated a "temporal response function" (TRF) following the approach that has previously been used for scalp potential during continuous narratives.[39,40] A TRF captures the impulse response from an "input" signal to an "output" signal (see STAR Methods). Here, the input signal is the gaze variation over time and the output signal is the scalp potentials. We find a positive anterior-central neural activity preceding fluctuations in gaze variation (100–300 ms) (Figure 2G). This differs substantially from the scalp potentials associated with gaze position (vertical in Figure S4B and horizontal in Figure S4C), blinks (Figure S4D), and saccades (Figure S4E). Here, scalp potentials are largely following the onset of these movements, which reflects well-known eye movement artifacts. In contrast, the anterior-central activity is predictive of subsequent changes in gaze variation.

Pupil size is affected by cognitive processing of the auditory narrative

As part of hypothesis H1, we also predicted a top-down effect of cortical processing of the narrative on pupil size. To test this, we measured the TRF of scalp potentials in response to pupil size fluctuations for each electrode during attentive listening to the narratives. We see a clear "response" (please note that we do

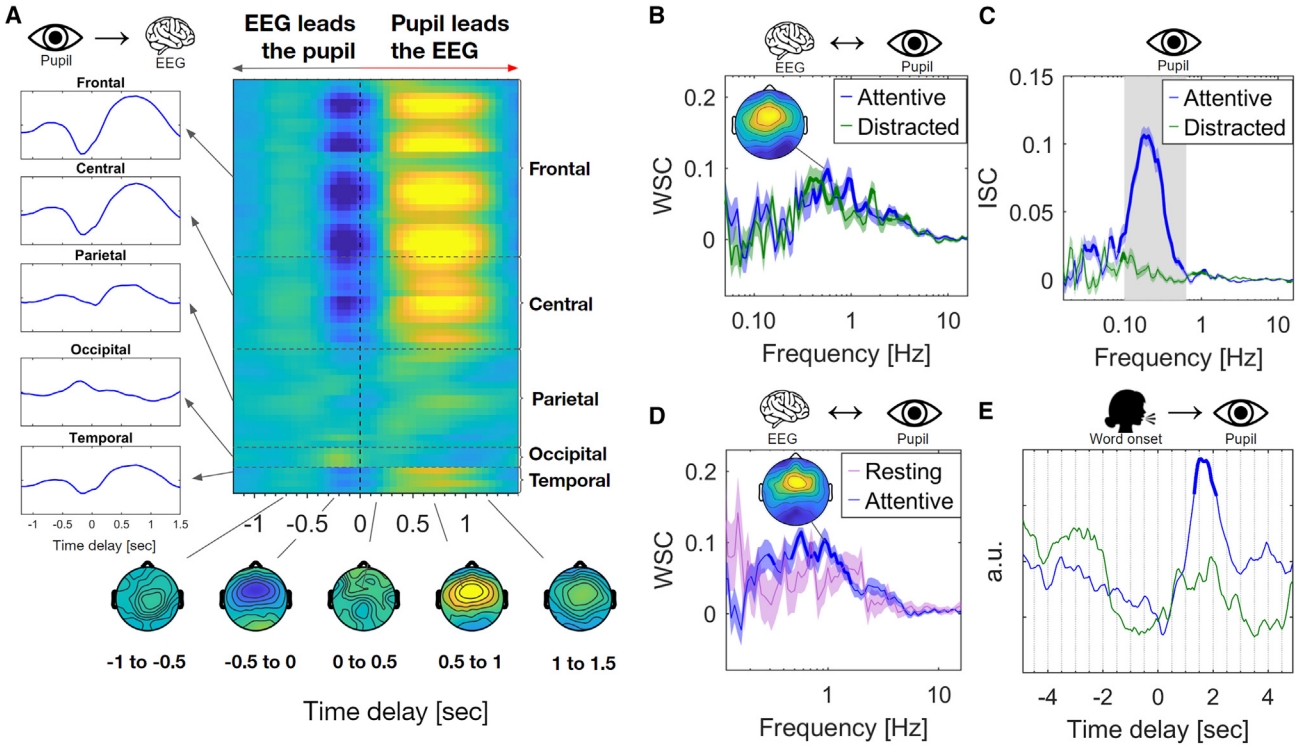


Figure 3. Pupil size during listening to auditory narrative

(A) TRFs using each EEG electrode as output signal and pupil size as input signal during the auditory narratives. Color indicates scalp potential fluctuations associated with pupil size (yellow and blue are positive and negative correlations, respectively). The dotted line indicates zero lag between the pupil size and scalp potentials (experiment 1, $n = 29$ subjects). The left images show the average TRFs for the frontal, central, parietal, occipital, and temporal brain areas.

(B) WSC between pupil size and scalp potentials during attentive and distracted listening to the narrative. Bold lines indicate significant WSC computed using t-statistics across subjects with cluster correction ($p < 0.01$, k = 101 frequency bins, $n = 29$ subjects).

(C) Frequency-resolved ISC computed by first band-pass filtering the pupil size at different frequencies and then computing ISC averaged over stories and subjects. Colored shading indicates SEM across $n = 29$ subjects. Bold lines indicate significant ISC and gray shading indicates a significant difference between attend and distracted conditions ($p < 0.01$, k = 101 frequency bins, $n = 29$ subjects). Significance of ISC is established in each band using the same method as in (B).

(D) WSC between pupil size and scalp potentials during attentive listening and resting state. Gray area indicates a significant difference between the resting and attentive conditions computed using t-statistics with cluster correction ($p < 0.01$, k = 101 frequency bins, $n = 29$ subjects).

(E) TRFs using pupil size as output signal and word onset as input. Bold line indicates a significant difference in TRF between conditions computed using shuffle statistics with false discovery rate (FDR) correction ($p < 0.01$, k = 160 time points, $n = 29$ subjects).

not mean to imply that the cortex is "responding" to changes in pupil size; we only use the terminology of TRF to stay consistent with the terminology that was used when linear systems modeling was introduced to EEG analysis[51,52]) at positive and negative time lags (Figure 3A), indicating that pupil fluctuations both precede and follow neural activity with no clear causal direction. The spatial distribution of this correlated brain activity has at least two distinct patterns on the scalp: an earlier anterior-central negativity (−500 to 0 ms), which is repeated positively (300 ms to 1 s).

To determine which timescale of fluctuations dominates this effect, we analyze the correlation between scalp potentials and pupil size but now resolved by frequency for any time delay (Figures 3B and 3D). We refer to this coherence measure as within-subject correlation (WSC).[34] Correlation is evident in a broad frequency range from 0.3 to 3 Hz ($N = 29$, k = 101, $p < 0.01$, bold lines in Figures 3B and 3D). The associated spatial patterns indicate that the faster activity has an anterior-central distribution (Figure 3B). Interestingly, when comparing the attentive versus distracted listening conditions, we see no difference in the strength of WSC. This suggests that the pupil-cortex coupling is largely unrelated to cognitive processing of the narrative and, instead, appears to be an endogenous link independent of an auditory stimulus.

The question remains, however, whether cognitive processing of the narrative drives autonomic activity reflected in the pupil. We test this by measuring ISC of the pupil and resolve this by frequency (Figure 3C). Significant ISC is present around the 0.1–0.5 Hz band while participants attentively listen to the narrative ($n = 29$, k = 101, $p < 0.01$ in 0.03–0.67 Hz and 0.93–1.15) but not when they are distracted, nor when they are attending to scrambled audio (Figure S6G). Considering that these narratives also entrain scalp potentials,[37] this suggests that the pupil is driven by cognitive processing of the narrative via a link of cortical activity acting on the pupil. Consistent with this, we find that the overall strength of ISC of pupil size of an individual

subject with the group is predictive of how well that individual recalls information about the narrative in the subsequent questionnaire (Figures S6E, $r(27) = 0.51$, $p = 4.8 \times 10^{-3}$, and S6F, $r(36) = 0.49$, $p = 1.7 \times 10^{-3}$). What is it in the narratives that could elicit pupil fluctuations?

First, we measured the response of pupil size to word onset using the same TRF formalism as before. We see a pupil dilation at a latency of 1.25–2.13 s ($n = 29$, $p < 0.01$, $k = 160$, bold line in Figure 3E), and this effect is reduced when the participants are distracted from the narrative (see raw traces of responses time locked to word onsets in Figures S9A–S9H). Could this be associated with sound volume fluctuations, which are known to elicit correlated fluctuations in scalp potentials?[41] To test this, we look at how the pupil size responds to the sound envelope, i.e., volume fluctuations of speech. We find that the TRF of the pupil size has non-zero responses at 1 s latencies, and this is true for both attended and unattended speech (Figures S6A and S6B, $n = 29$, $k = 160$, $p < 0.01$, bold lines). In short, pupil responses to word onsets are modulated by attention, yet responses to lower-level sound volume fluctuations are not.

Another aspect that could drive pupil fluctuations is the complexity of the speech and associated listening effort.[13] As a measure of lexical difficulty, we measured the bit rate for each word (see STAR Methods). We jointly estimated the TRF of the pupil size to both word offset and bit rate so as to capture an incremental contribution from lexical difficulty. We found a peak in the TRF around 1–1.5 s (Figure S6C).

One caveat of our results on pupil size is that saccades and eye blinks are known to cause pupil size fluctuations.[43,44] Indeed, we confirm the effect of saccade and blinks on pupil size (Figure S1). Therefore, in the preceding analysis we removed any linear effect due to saccades and blinks from the pupil size measurements (see Figure S1). Therefore, none of the effects in Figure 3 can be attributed to the saccade and pupil effect reported in Figure 2 except for the ISC effect, which can capture non-linear pupil responses that are similar across subjects.

Heart rate is affected by cognitive processing of the auditory narrative and correlates with scalp potentials

Hypothesis H1 of a top-down effect of cognition on autonomic function also predicted a correlation of scalp potentials with heart rate. We again measure TRFs, now with the instantaneous heart rate (see STAR Methods) as input and scalp potentials as output. We find non-zero responses to heart rate, mostly leading the neural signal (Figure 4A). Interestingly, we see again an anterior-central component in the scalp potentials that is anticipated by the heart rate by around 1–1.4 s. The same pattern is present for responses in the distracted and rest conditions, suggesting that this heart-cortex link is endogenous.

As predicted by hypothesis H1, we also find that the narrative synchronizes heart rate between subjects (Figure 4B). The effect is stronger in the attentive listening condition ($n = 29$, $k = 101$, $p < 0.01$ in 0.02–0.12 Hz), with an intermediate value in the scrambled audio condition (Figure S7F, $n = 38$, $k = 101$, $p < 0.01$ in 0.03–0.05 and 0.12–0.20 Hz). This replicates our previous findings on video[33] and confirms that cognitive processing of the narrative drives some of these slow-frequency (0.1 Hz) heart rate fluctuations. A common measure of the effect of cognition on heart rate is heart rate variability, which, when resolved by frequency, is simply the power spectrum of the instantaneous heart rate (Figure 4C). The well-established high-frequency (HF-HRV) and low-frequency (LF-HRV) components of heart rate variability[53] are quite evident here, but we see no effect of attention. We note that the high-frequency component corresponds to rhythmic breathing, which modulates heart rate at the breathing frequency.[54] In our data, this is most evident in the WSC between respiration and heart rate (Figure 4D). We see the respiratory sinus arrhythmia as a peak at 0.3 Hz. However, the synchronization of heart rate across subjects is not the result of synchronized breathing, as there is no reliable ISC for breathing during the narratives (Figures S7A and S7B). This is also evident in the instantaneous heart rate signal aligned to peak inhalation (Figure 4E).

Respiratory volume has recently been linked to arousal, and it has been shown to correlate with the global signal in fMRI at ultra-slow fluctuations.[55] In our data, we find a synchronization of respiration volume across subjects, with a modulation by attention around 0.03 Hz (Figures 4F, $n = 29$, $k = 101$, $p < 0.01$ in 0.03–0.04 and 0.07–0.08 Hz, and S7C). However, we do not find any coupling between scalp potentials and respiratory volume at these very slow fluctuations (Figure S7D), but we cannot rule out limitations of the EEG equipment in this low-frequency range.

Bottom-up effect of heart rate control but not pupil control

Hypothesis H2 posits a bottom-up effect of autonomic function. By this, we mean that brain nuclei controlling, e.g., pupil size may also affect heart rate, or vice versa, in the absence of a top-down cognitive drive. To test this, we analyzed data collected during controlled breathing and controlled luminance, with both conditions in silence, and compared these to resting state. As the controlled interventions were time aligned across subjects (see STAR Methods), we can again use ISC to capture reliable responses to these interventions. Assuming that effects are the same across subjects, this approach can capture linear as well as non-linear or delayed responses to the interventions.

In the controlled luminance condition, the brightness of the screen was modulated at a rate of 0.1 Hz. This elicited an obvious synchronization of pupil size across subjects (Figure 5A, $t(28) = 31.42$, $p = 3.24 \times 10^{-23}$), as well as a synchronization of eye movements (gaze position, $t(28) = 3.58$, $p = 1.28 \times 10^{-3}$, and saccade rate, $t(28) = 3.94$, $p = 4.93 \times 10^{-4}$). Note that there was no fixation cross, so it is conceivable that the change in brightness may have prompted subjects to move their gaze in a predictable way and synchronization in the EEG ($t(28) = 3.15$, $p = 3.91 \times 10^{-3}$) is possibly just the result of eye movement artifacts. Importantly, we find no effects on heart rate or breathing.

Synchronized breathing, on the other hand, had broad-ranging effects (Figure 5B), synchronizing not only respiration ($t(28) = 4.60$, $p = 8.21 \times 10^{-5}$) but heart rate ($t(28) = 4.02$, $p = 3.95 \times 10^{-4}$) and pupil size ($t(27) = 2.91$, $p = 7.13 \times 10^{-3}$). We can directly compare the bottom-up effect of controlled breathing and luminance by computing the WSC between heart rate and pupil size (Figures 5D and 5E) and compare that to when participants are at rest. Here, we see a significant peak in the WSC at 0.1 Hz of the controlled breathing ($n = 29$, $k = 61$,

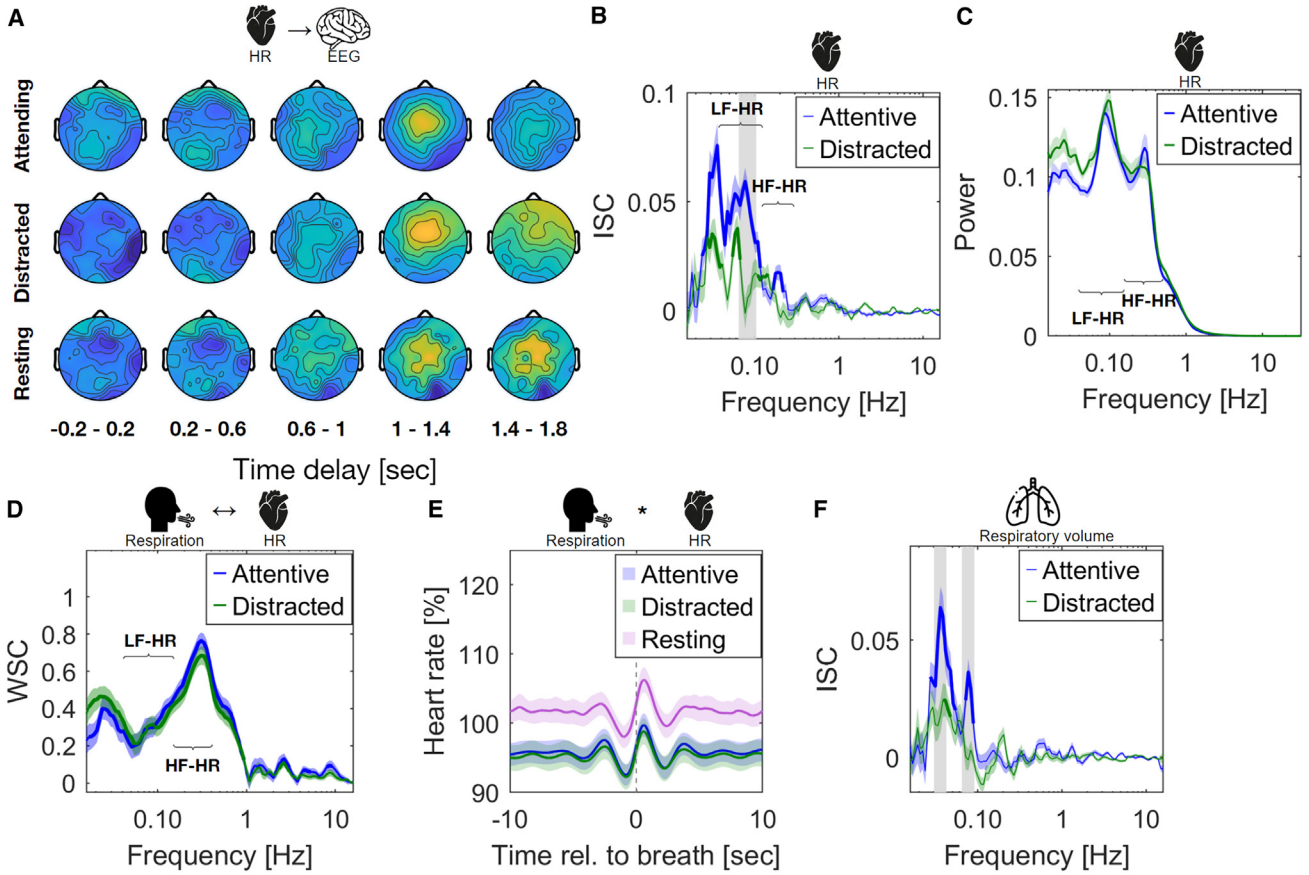


Figure 4. Heart rate and breathing during listening to auditory narratives

(A) Topographical plot of TRF between heart rate and scalp potentials similar to Figure 2A. Positive lag indicates that heart rate fluctuations lead scalp potential fluctuations (n = 29).

(B) Frequency-resolved ISC of heart rate. Colored shaded area indicates SEM across n = 29 participants. Gray shaded area indicates frequency bands with significant differences between attend and distracted listening conditions. Significance of ISC is established in each band using t-statistics, corrected for multiple comparisons using one-dimensional cluster statistics ($p < 0.01$, k = 101 frequency bins, n = 29). LF-HR is the low-frequency heart rate region and HF-HR is the high-frequency heart rate range.

(C) Power spectrum of heart rate computed as the normalized power (dividing the power in each frequency bin by the aggregate power of the attentive and distracted conditions). Significance between the attentive and distracted conditions was compared as in (B) but using the Wilcoxon rank-sum test instead of the t-statistics.

(D) WSC between respiration and heart rate (statistics and error bars computed like in B).

(E) Breathing-evoked heart rate fluctuations relative to baseline (10 s average before each story was played) (reverse correlation).

(F) Frequency-resolved ISC for respiratory volume. Error bars and statistics were computed as in (B).

$p < 0.01$, bold line and gray shading in Figure 5E). This is the same frequency at which participants were asked to breathe. However, this correlation is not present for the controlled luminance condition (with luminance modulated at the same frequency). This suggests that while brain nuclei controlling pupil size have a fairly narrow effect, the effect of modulating heart rate using respiration seems to affect a number of autonomic signals. Consistent with this, we find that the correlation of heart rate with pupil size remains when participants were at rest with a peak at 0.3 Hz (n = 29, k = 101, $p < 0.01$, bold line in Figure 5E), corresponding to the natural breathing rhythm.

When we listen to auditory narratives and breathe at our natural breathing rate, the pupil heart rate link remains (n = 29, k = 101, $p < 0.01$, bold line in Figure 5F). However, we do not see any consistent significant effect of attention, suggesting an endogenous link. We also analyzed the relative time delay between these coherent fluctuations using the formalism of TRF. In this case, we model the linear response of the pupil size onto heart rate, and vice versa (n = 29, k = 160, $p < 0.01$ in bold line in Figure 5C). In both instances, we obtain non-zero responses with fluctuations in heart rate that are ahead of fluctuations in pupil size.

Bottom-up effect of eye movements on heart rate

The other peripheral signal we explored was eye movements. A bottom-up effect following hypothesis H2 might predict that eye movement has an effect on heart rate. To test this, we asked subjects to follow dots on the screen that jumped in position every 2

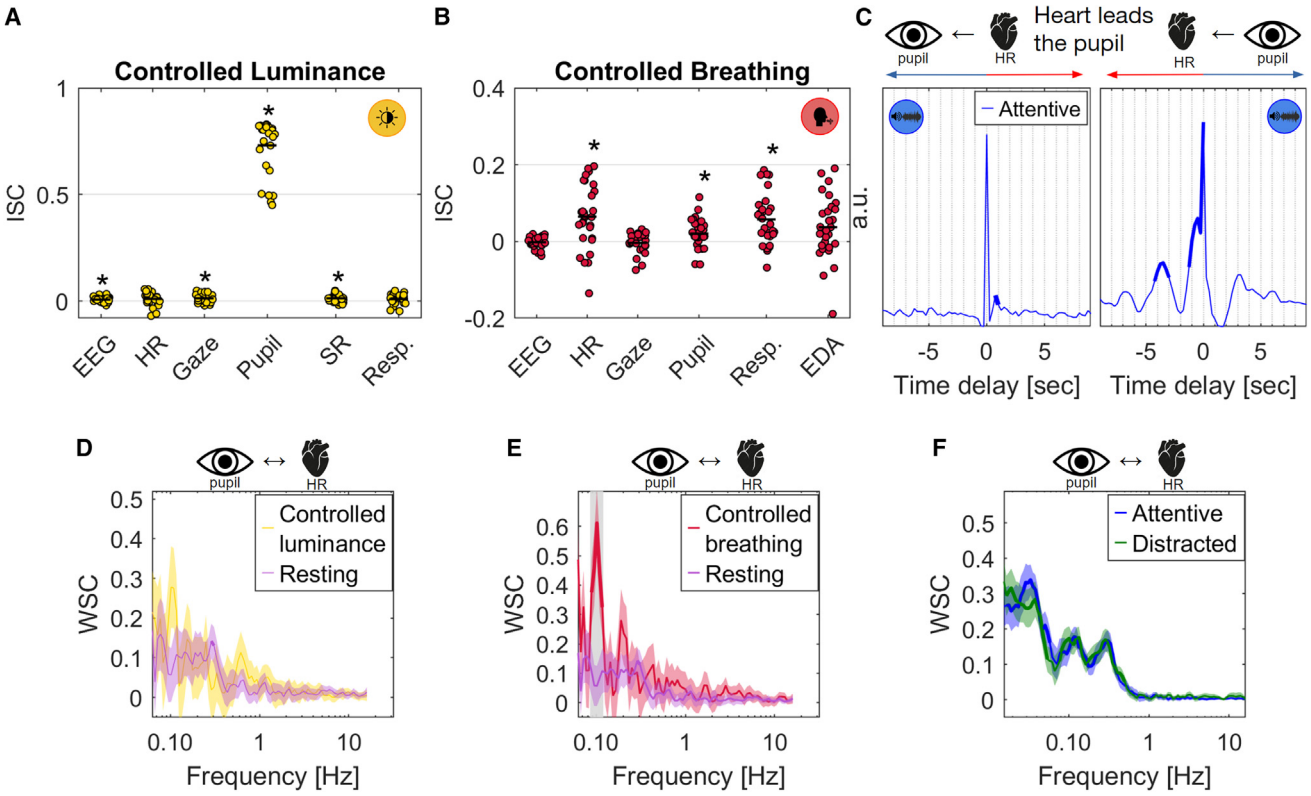


Figure 5. Cross-modal effects of controlled luminance and controlled breathing

(A) ISC during the controlled luminance condition. The asterisk (*) indicates that the ISC values are significantly different from zero using a t test ($p < 0.01$, $n = 29$).

(B) ISC during the controlled breathing condition. Significance was computed as in (A).

(C) TRF between heart rate and pupil size during attentive listening. Left image models the linear response of pupil size (as output) to heart rate (as input). The right image models the reverse. Non-zero response (for positive lags indicated using a blue arrow, and for negative lags indicated using a red arrow) indicates that fluctuations of heart rate are ahead of fluctuations in pupil size. Significance was obtained using shuffle statistics ($p < 0.01$, k = 160 time points, $n = 29$).

(D) WSC between pupil size and heart rate during controlled luminance condition and resting state. Colored shaded area indicates the SEM and bold line indicates the WSC is significantly different from 0 ($n = 29$, k = 101 frequency bins, $p < 0.01$) (see STAR Methods).

(E) Same as (D) but for controlled breathing condition compared to resting state.

(F) Same as (E) but for attentive and distracted listening conditions.

s. Measuring ISC shows entrainment for a number of signals (Figure 6A). Eye movements obviously synchronize saccade rate ($t(28) = 11.49$, $p = 4.15 \times 10^{-12}$) and pupil size ($t(28) = 12.24$, $p = 9.39 \times 10^{-13}$) but also entrain EEG ($t(28) = 11.93$, $p = 1.69 \times 10^{-12}$), presumably due to eye movement artifacts on the scalp potentials. However, the significant ISC we observe for heart rate cannot be trivially explained ($t(28) = 6.06$, $p = 1.57 \times 10^{-6}$). When resolved by frequency, we see a clear synchronization at the 0.5 Hz rate of saccades ($n = 29$, k = 61, $p < 0.01$ in shaded area in Figure 6C), scalp potentials ($n = 29$, k = 61, $p < 0.01$ in shaded area in Figure 6B), pupil size ($n = 29$, k = 61, $p < 0.01$ in shaded area in Figure 6D), and lastly, heart rate ($n = 29$, k = 61, $p < 0.01$ in shaded area in Figure 6E), suggesting a bottom-up effect of saccades on individual heart beats. A 0.5 Hz oscillation in heart rate reflects a slow beat followed by a fast beat, in synchrony with the rhythmic saccades. In the second experiment, we repeated this intervention with different rhythms of saccades (0.2, 0.5, 1, and 2 Hz) but only found significant synchronization for 0.5 Hz (Figures S8A–S8D). This suggests that saccades have an instantaneous effect on individual heart beats that manifests most clearly when saccades follow the natural rhythm of the heart (Figure 4C).

DISCUSSION

We set out to test whether pupil size and eye movements are modulated by listening to an auditory narrative. We had seen this previously for viewing of audiovisual narratives,[34] but it was not obvious that it should replicate in the absence of a visual stimulus. We hypothesized that these signals are peripheral markers of arousal and, as such, should be similarly driven by top-down effects of cognition. We confirmed this here for pupil size and heart rate and for saccade initiation and blinks.

Surprisingly, we found that slow variations of gaze position on a timescale of 10 s were driven by the narrative. Specific moments of high and low variation during a narrative were similar across listeners, suggesting that the content of the narrative affected gaze variation. We draw the same conclusions for fluctuations of pupil size and heart rate: ISC at a similar timescale of 10 s was modulated by attention, and it was predictive of

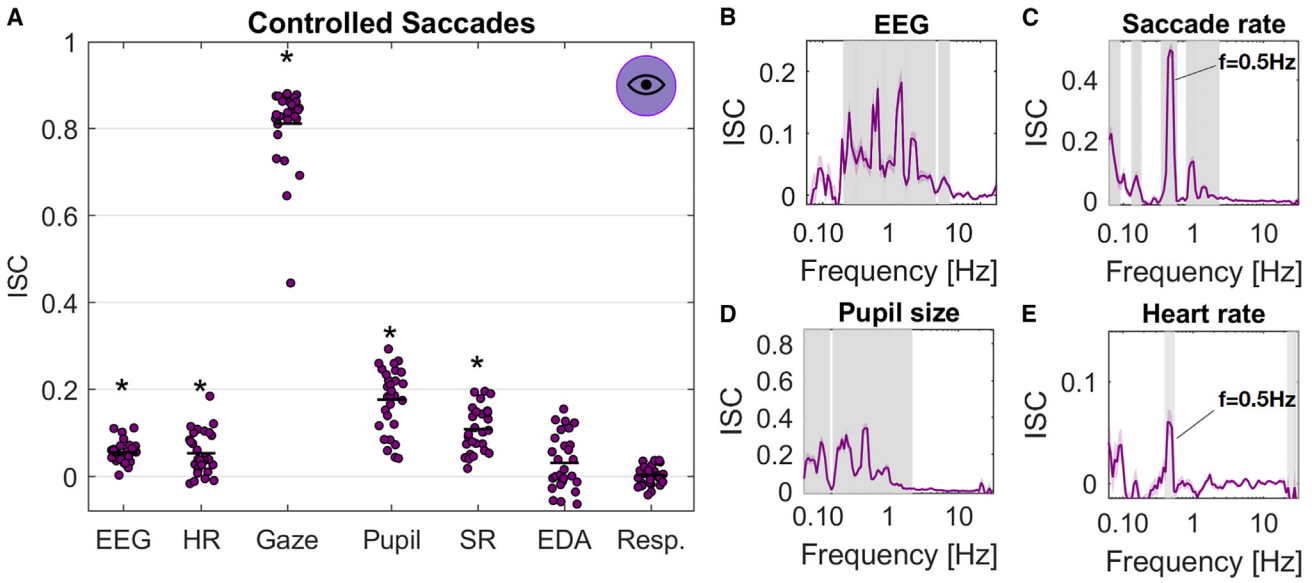


Figure 6. Rhythmic saccades modulate heart rate

(A) ISC of various signal modalities during the controlled saccade experiment. Participants followed a dot jumping on the screen in a 9 point grid every 2 s. The asterisk (*) indicates that the ISC values are significantly different from zero using a t test ($p < 0.01$, $n = 29$).

(B) Frequency-resolved ISC for EEG during the controlled saccade experiment. Line indicates the average across $n = 29$ participants, and the color shaded area indicates SEM for $n = 29$. Significance of ISC was computed using t-statistics in each frequency band, corrected for multiple comparisons using one-dimensional cluster statistics (gray shaded area indicates $p < 0.01$, $n = 29$, k = 61 frequency bins).

(C) Same as (B) but for saccade rate.

(D) Same as (B) but for pupil size.

(E) Same as (B) but for heart rate.

memory of the auditory narrative. Timing of blinks and saccades also entrained to the narrative at similarly slow timescales. Together, these results suggest that cognitive processing of the narrative drives autonomic function, perhaps reflecting varying levels of arousal. The effect of cognition and autonomic arousal is consistent with the notion of allostasis, which is the efficient regulation required to prepare the body for future demands.[56] In this view, cognitive processing of the narrative may be driving autonomic function to prepare the body to act.

We also set out to quantify bottom-up effects, whereby a low-level drive of these peripheral signals has an effect on the other autonomic signals. While there is some overlap in brain nuclei known to control pupil size, heart rate, and eye movements, the control networks for the corresponding effector muscles are quite distinct, and bottom-up afferent connections are not as clearly established. Breathing has a well-established effect on heart rate[54] and is subject to a combination of central control and a peripheral effect such as the baroreflex.[21] But based on the neuroanatomical literature alone, it was not obvious that breathing would affect pupil size, yet this is what we found here. Another well-established, low-level effect is pupillary light reflex.[1] Luminance fluctuations might have affected arousal and thus driven heart rate. We did not see this here. Instead, we found that heart rate fluctuations preceded pupil fluctuations, and in fact, they preceded slow cortical potential fluctuations. This suggests that cardiac control has a broad bottom-up effect on arousal, which pupil control does not. Indeed, there is a long list of studies arguing for a bottom-up effect of cardiac function of the brain,[27,28,57] with the effect often attributed to arousal.[58,59] In this context, it may be worth noting that there are global fluctuations in neural activity across the brain that are coherent with the fMRI signal at and below 0.1 Hz.[60] In rodents, spontaneous global fluctuations in neural activity at 0.1 Hz are synchronous with vascular dilatation in the brain.[61] These are controlled by midbrain structures (e.g., rostral ventrolateral medulla), which also regulate sympathetic drive of the heart.[62] This may provide a viable physiological mechanism for the bottom-up sympathetic drive of slow fluctuations observed here.

We found that scalp potentials correlate with heart rate, pupil size, and gaze variation. The spatial distribution for all three has an anterior-central distribution, suggesting a common cortical source. The anterior-central distribution is reminiscent of the "error-related negativity"[63] and is consistent with a source in the anterior cingulate (ACC).[64] A recent integrative theory of the LC-norepinephrine function argues that that the ACC plays a role in optimizing the utility between short-timescale LC activity (phasic pupillary responses) and long timescales (tonic pupillary responses), which are associated with disengagement from task demand. The slow fluctuations we find synchronizing between people could be related to these switches between internal and external attention. This interpretation is consistent with the finding that ISC of scalp potentials during audiovisual narratives correlates positively with fMRI activation of the ACC, among other brain areas.[65] ISC of fMRI activity in the ACC is also predictive of performance in a quiz on the material presented in educational videos,[66] similar to what we found previously with EEG.[67]

The power of LF-HRV, which is often associated with sympathetic arousal,[53] correlates positively with fMRI activity in the ACC,[68] and heart rate itself correlates with fMRI in midline brain structures including the ACC.[25] Indeed, the ACC is thought to be one of the cortical areas regulating autonomic function.[27] Additionally, activity in the ACC is modulated by activity in the LC,[69] which is tightly coupled with pupil size and often taken as an indicator of central arousal.[5] While we emphasized the ACC here, there may be other sources of brain activity with the same frontocentral scalp distribution. Future work with magnetoencephalography (MEG) or fMRI may be able to address our proposal that ACC activity correlates with slow fluctuations of autonomic signals during attentive listening to narratives.

We found that the speech tracking phenomenon observed for scalp potentials[38–41] carries over to the pupil that responds to the sound envelope, word onset, and lexical difficulty and is modulated by attention. We also found that saccade initiation and blink onset were driven by the narrative in the absence of a visual stimulus. This may open up opportunities for new research on speech processing[39] and auditory attention decoding[70] using eye tracking and, in particular, the pupil.

In analyzing eye movements, we were motivated by the recent proposal that saccade initiation and pupil dilation are subject to a common central drive.[15] Saccades are driven by activity in the superior colliculus, which receives input from the LC, and that, in turn, is known to affect heart rate.[4] Given these overlapping brain nuclei, there was a possibility that cognitive processing of the narrative might have a top-down effect on saccade initiation. Conversely, the overlapping control nuclei may mediate a bottom-up effect of saccades on heart rate, or heart rate on saccades, in the absence of a cognitive effect. Indeed, we found a top-down effect of cognition on saccade initiation. Our observation that saccades and blinks entrain to the narrative during attentive listening is consistent with the recent reports that gaze position, saccades, and blinks correlated with low-level features of continuous speech.[50,71,72] We also found a bottom-up effect of saccades on individual heart beats. This is consistent with a previous observation that saccades tend to align in time with the cardiac cycle during visual search.[73] A link between saccades and heart rate has also been postulated in the context of "eye movement desensitization and reprocessing therapy." This intervention asks patients to recall traumatic experiences while eliciting rhythmic left/right eye movements (every 0.5 s). There are some reports that this intervention lowers heart rate during the procedure.[74,75] We find that heart beats entrain to saccades only if they are repeated every 2 s, so the relationship to this clinical intervention is unclear.

An unexpected result is that gaze variation, and not saccades, was larger during attentive listening and was driven by the narrative. This suggests that low-level control of eye movements was not altered but simply that our gaze canvases a wider area of the visual field when listening to a story. This is consistent with a recent study showing that gaze variation (about a fixation cross) during prestimulus baseline correlated negatively with alpha power in the EEG.[76] Both alpha power and gaze variation were predictive of subsequent memory of the upcoming visual scene. Alpha is generally considered a measure of cortical inhibition.[77] These results are therefore consistent with an interpretation of gaze variation as a measure of arousal. Interestingly, we found that anterior-central activity was predictive of subsequent changes in gaze variation, suggesting a top-down effect of the processing of the narrative on gaze variation. While eye movements during listening to a narrative may serve an ecological function, in particular in a social setting, in the present setting, eye movements are not required for the task and therefore are incidental to the task. Movements unrelated to a task have recently been shown to dominate neural activity in cortical and subcortical brain areas in rodents.[78,79] But the relationship of incidental behaviors to arousal has been unclear thus far. Here, we have identified a clear association of a readily observable incidental behavior with traditional markers of autonomic arousal.

Limitations of the study

Some level of synchronization between subjects was also observed for the heart rate when participants attended to modulated noise and, to a lesser degree, when they heard the narrative but were engaged in a mental arithmetic task. This suggests that low-level features of the audio can drive physiological responses, perhaps by modulating arousal or effort.

To rule out effects of saccade initiation and blinks on pupil size,[43,44] we removed linear effects between these signals prior to analyzing pupil size. It is possible that we did not adequately account for these interactions. On the other hand, these interactions are consistent with our hypothesis, and prior theory, that links saccade initiation and pupil to a common physiological drive related to arousal.[15]

Eye movement and blinks cause obvious artifacts in the EEG. While they can be removed fairly reliably using linear regression,[80] it is possible that non-linear artifacts remain. These artifacts are relatively fast (<100 ms), whereas most of the findings in this study are at lower frequency that are likely unaffected. It is important to acknowledge that the fluctuations driven by the narrative for both heart rate and pupil are relatively slow (~10 s) as compared to the link we find between these peripheral signals and the brain (~1 s). Interestingly, we do not detect pupil ISC in the faster 1 Hz band, which we previously observed with audiovisual stimuli,[34] suggesting that these faster pupil fluctuations were due to the dynamics of the visual stimulus and/or saccades.

An alternative explanation of our results at these slow timescales is that increased heart rate (and breathing) increases blood oxygenation. However, while the breathing cycle itself had no effect, it is possible that heart rate is affected by breathing volume. Either way, respiration volume appears to be another signal reflecting the top-down effect of cognition on arousal, similar to heart rate and gaze variation.

When measuring correlation of neural activity with the peripheral signals, we focused on fluctuations of the raw potentials and not their oscillatory power. The motivation for this is the extensive literature on "evoked responses"' during continuous natural stimuli, such as speech and videos.[39,51,65,81,82] Compared to conventional evoked responses, the timescale of fluctuations reported here is relatively slow, likely due to the slow timescale of the peripheral signals we are correlating with. Nevertheless, the question arises as to what these slow

fluctuations represent. The fluctuations are comparable to "slow scalp potentials," which have been reported in the context of cognition and action in early EEG work.[83] Evoked responses are generally thought to reflect large-scale waves of synaptic activity (as opposed to neuronal firing), but the physiological origins of slow fluctuations are not well understood. A more recent report suggests that fluctuations in the 0.1 Hz range correspond to hemodynamic activity in the brain,[84] although it is worth noting that the fluctuations we observed in the EEG are not as coherent as those reported there.

Conclusions

By simultaneously recording brain activity and a variety of peripheral signals, we investigated how the brain and body interact during natural behaviors. Our overall conclusion is that the coupling of the central and autonomic nervous system is bidirectional and fairly independent of task performance. Cognitive processing of a narrative modulates arousal and thus arousal-related physiological signals. As a result, the endogenous dynamic of the entire body is engaged when listening to stories. We conjecture that the endogenous dynamic is entrained by the narrative at a preferred timescale of 0.1 Hz. In that sense, one may say that the brain-body dynamic "resonates" with the narrative. The present results add to the evidence that arousal mediates the interaction between cognition and autonomic function. Many of the interactions observed here have been studied in isolation within neuroscience, physiology, and psychology and may have missed the whole-body dynamic reported here. We also see a relevance for research on auditory processing, cognitive effects of arousal, and slow signal fluctuations observed in the brain.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Auditory stimuli
  - Procedure
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Alignment of multimodal recordings
  - Recording and preprocessing of gaze position, gaze variation and pupil size
  - Recording and preprocessing of EEG
  - Recording and preprocessing of ECG
  - Recording and preprocessing of respiration and respiratory volume
  - Inter-subject correlation (ISC)
  - Frequency resolved inter-subject correlation (synchronization spectrum)
  - Frequency resolved within-subject correlation (WSC)
  - Cluster statistics for contrast between conditions in frequency resolved correlation
  - Cluster statistics for significance of frequency resolved correlation
  - Temporal response function (TRF)
  - Significance of temporal response function
  - Speech features

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114081

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation through grant no. DRL-2201835. We want to thank Alexandria Hoang for her help with collecting some of these data. We would like to thank Vinay S. Raghavan for alerting us about the potential confound of blinks on pupil size.

AUTHOR CONTRIBUTIONS

J.M. and L.C.P. contributed to the conception and design of the study. J.M. contributed to the acquisition of data. J.M. and L.C.P. contributed to the analysis of data. J.M. and L.C.P. contributed to the writing of the manuscript and/or the figures.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 1, 2023
Revised: November 25, 2023
Accepted: March 24, 2024
Published: April 6, 2024

REFERENCES

1. Ellis, C.J. (1981). The pupillary light reflex in normal subjects. Br. J. Ophthalmol. *65*, 754–759.
2. Turnbull, P.R.K., Irani, N., Lim, N., and Phillips, J.R. (2017). Origins of Pupillary Hippus in the Autonomic Nervous System. Invest. Ophthalmol. Vis. Sci. *58*, 197–203.
3. Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. J. Cogn. *1*, 16.
4. Samuels, E.R., and Szabadi, E. (2008). Functional Neuroanatomy of the Noradrenergic Locus Coeruleus: Its Roles in the Regulation of Arousal and Autonomic Function Part I: Principles of Functional Organisation. Curr. Neuropharmacol. *6*, 235–253.
5. Costa, V.D., and Rudebeck, P.H. (2016). More than Meets the Eye: the Relationship between Pupil Size and Locus Coeruleus Activity. Neuron *89*, 8–10.
6. Joshi, S., and Gold, J.I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. Trends Cognit. Sci. *24*, 466–480.
7. Hess, E.H., and Polt, J.M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. Science *143*, 1190–1192.
8. Bradshaw, J. (1967). Pupil Size as a Measure of Arousal during Information Processing. Nature *216*, 515–516.
9. Granholm, E., Asarnow, R.F., Sarkin, A.J., and Dykes, K.L. (1996). Pupillary responses index cognitive resource limitations. Psychophysiology *33*, 457–461.
10. Hess, E.H., and Polt, J.M. (1960). Pupil Size as Related to Interest Value of Visual Stimuli. Science *132*, 349–350.

11. Bradley, M.M., Miccoli, L., Escrig, M.A., and Lang, P.J. (2008). The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology *45*, 602–607.
12. Partala, T., and Surakka, V. (2003). Pupil size variation as an indication of affective processing. Int. J. Hum. Comput. Stud. *59*, 185–198.
13. van der Wel, P., and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. Psychon. Bull. Rev. *25*, 2005–2015.
14. Grueschow, M., Kleim, B., and Ruff, C.C. (2020). Role of the locus coeruleus arousal system in cognitive control. J. Neuroendocrinol. *32*, e12890.
15. Burlingham, C.S., Mirbagheri, S., and Heeger, D.J. (2022). A unified model of the task-evoked pupil response. Sci. Adv. *8*, eabi9979.
16. Quirins, M., Marois, C., Valente, M., Seassau, M., Weiss, N., El Karoui, I., Hochmann, J.R., and Naccache, L. (2018). Conscious processing of auditory regularities induces a pupil dilation. Sci. Rep. *8*, 14819.
17. Konishi, M., Brown, K., Battaglini, L., and Smallwood, J. (2017). When attention wanders: Pupillometric signatures of fluctuations in external attention. Cognition *168*, 16–26.
18. Wang, C.-A., and Munoz, D.P. (2015). A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. Curr. Opin. Neurobiol. *33*, 134–140.
19. Aston-Jones, G., and Cohen, J.D. (2005). AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. Annu. Rev. Neurosci. *28*, 403–450.
20. Mendelowitz, D. (1999). Advances in Parasympathetic Control of Heart Rate and Cardiac Function. Physiology *14*, 155–161.
21. Berntson, G.G., Cacioppo, J.T., and Quigley, K.S. (1993). Respiratory sinus arrhythmia: Autonomic origins, physiological mechanisms, and psychophysiological implications. Psychophysiology *30*, 183–196.
22. Critchley, H.D., Corfield, D.R., Chandler, M.P., Mathias, C.J., and Dolan, R.J. (2000). Cerebral correlates of autonomic cardiovascular arousal: a functional neuroimaging investigation in humans. J. Physiol. *523*, 259–270.
23. Wang, C.-A., Baird, T., Huang, J., Coutinho, J.D., Brien, D.C., and Munoz, D.P. (2018). Arousal Effects on Pupil Size, Heart Rate, and Skin Conductance in an Emotional Face Task. Front. Neurol. *9*, 1029.
24. Raimondo, F., Rohaut, B., Demertzi, A., Valente, M., Engemann, D.A., Salti, M., Fernandez Slezak, D., Naccache, L., and Sitt, J.D. (2017). Brain–heart interactions reveal consciousness in noncommunicating patients. Ann. Neurol. *82*, 578–591.
25. Valenza, G., Sclocco, R., Duggento, A., Passamonti, L., Napadow, V., Barbieri, R., and Toschi, N. (2019). The central autonomic network at rest: Uncovering functional MRI correlates of time-varying autonomic outflow. Neuroimage *197*, 383–390.
26. Shaffer, F., and Ginsberg, J.P. (2017). An Overview of Heart Rate Variability Metrics and Norms. Front. Public Health *5*, 258.
27. Park, G., and Thayer, J.F. (2014). From the heart to the mind: cardiac vagal tone modulates top-down and bottom-up visual perception and attention to emotional stimuli. Front. Psychol. *5*, 278.
28. Mather, M., and Thayer, J. (2018). How heart rate variability affects emotion regulation brain networks. Curr. Opin. Behav. Sci. *19*, 98–104.
29. Thayer, J.F., Hansen, A.L., Saus-Rose, E., and Johnsen, B.H. (2009). Heart Rate Variability, Prefrontal Neural Function, and Cognitive Performance: The Neurovisceral Integration Perspective on Self-regulation, Adaptation, and Health. Ann. Behav. Med. *37*, 141–153.
30. McCall, C., Hildebrandt, L.K., Bornemann, B., and Singer, T. (2015). Physiophenomenology in retrospect: Memory reliably reflects physiological arousal during a prior threatening experience. Conscious. Cognit. *38*, 60–70.
31. Hansen, A.L., Johnsen, B.H., and Thayer, J.F. (2003). Vagal influence on working memory and attention. Int. J. Psychophysiol. *48*, 263–274.
32. Appelhans, B.M., and Luecken, L.J. (2006). Heart Rate Variability as an Index of Regulated Emotional Responding. Rev. Gen. Psychol. *10*, 229–240.
33. Pérez, P., Madsen, J., Banellis, L., Türker, B., Raimondo, F., Perlbarg, V., Valente, M., Niérat, M.C., Puybasset, L., Naccache, L., et al. (2021). Conscious processing of narrative stimuli synchronizes heart rate between individuals. Cell Rep. *36*, 109692.
34. Madsen, J., and Parra, L.C. (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. PNAS Nexus *1*, pgac020.
35. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject Synchronization of Cortical Activity During Natural Vision. Science *303*, 1634–1640.
36. Simony, E., Honey, C.J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., and Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. *7*, 12141.
37. Cohen, S.S., and Parra, L.C. (2016). Memorable Audiovisual Narratives Synchronize Sensory and Supramodal Neural Responses. eNeuro *3*. 0203-0216.2016.
38. Ding, N., and Simon, J.Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. Front. Hum. Neurosci. *8*, 311.
39. Di Liberto, G.M., O'Sullivan, J.A., and Lalor, E.C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. Curr. Biol. *25*, 2457–2465.
40. Broderick, M.P., Anderson, A.J., Di Liberto, G.M., Crosse, M.J., and Lalor, E.C. (2018). Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. Curr. Biol. *28*, 803–809.e3.
41. Zion Golumbic, E.M., Ding, N., Bickel, S., Lakatos, P., Schevon, C.A., McKhann, G.M., Goodman, R.R., Emerson, R., Mehta, A.D., Simon, J.Z., et al. (2013). Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a "Cocktail Party". Neuron *77*, 980–991.
42. Ki, J.J., Kelly, S.P., and Parra, L.C. (2016). Attention Strongly Modulates Reliability of Neural Responses to Naturalistic Narrative Stimuli. J. Neurosci. *36*, 3092–3101.
43. Knapen, T., de Gee, J.W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., and Theeuwes, J. (2016). Cognitive and Ocular Factors Jointly Determine Pupil Responses under Equiluminance. PLoS One *11*, e0155574.
44. Yoo, K., Ahn, J., and Lee, S.-H. (2021). The confounding effects of eye blinking on pupillometry, and their remedy. PLoS One *16*, e0261463.
45. Perry, S., Khovanova, N.A., and Khovanov, I.A. (2019). Control of heart rate through guided high-rate breathing. Sci. Rep. *9*, 1545.
46. Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., and Tolias, A.S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. Neuron *84*, 355–362.
47. Lloyd, B., Voogd, L. D. de, Mäki-Marttunen, V., and Nieuwenhuis, S. (2022). Assessing pupil size as an index of activation of subcortical ascending arousal system nuclei during rest, . Preprint at. https://doi.org/10.1101/2022.11.04.514984.
48. Larsen, R.S., and Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. Front. Neural Circ. *12*, 21.
49. Madsen, J., Júlio, S.U., Gucik, P.J., Steinberg, R., and Parra, L.C. (2021). Synchronized eye movements predict test scores in online video education. Proc. Natl. Acad. Sci. USA *118*, e2016980118.
50. Holtze, B., Rosenkranz, M., Bleichner, M.G., Jaeger, M., and Debener, S. (2022). Eye-blink patterns reflect attention to continuous speech, . Preprint at PsyArXiv. https://doi.org/10.31234/osf.io/n86yp.
51. Lalor, E.C., Pearlmutter, B.A., Reilly, R.B., McDarby, G., and Foxe, J.J. (2006). The VESPA: A method for the rapid estimation of a visual evoked potential. Neuroimage *32*, 1549–1561.
52. Crosse, M.J., Di Liberto, G.M., Bednar, A., and Lalor, E.C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. Front. Hum. Neurosci. *10*, 604.
53. Reyes del Paso, G.A., Langewitz, W., Mulder, L.J.M., van Roon, A., and Duschek, S. (2013). The utility of low frequency heart rate variability as

an index of sympathetic cardiac tone: A review with emphasis on a reanalysis of previous studies. Psychophysiology *50*, 477–487.

54. Hirsch, J.A., and Bishop, B. (1981). Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. Am. J. Physiol. *241*, H620–H629.
55. Raut, R.V., Snyder, A.Z., Mitra, A., Yellin, D., Fujii, N., Malach, R., and Raichle, M.E. (2021). Global waves synchronize the brain's functional systems with fluctuating arousal. Sci. Adv. *7*, eabf2709.
56. Sterling, P. (2012). Allostasis: A model of predictive regulation. Physiol. Behav. *106*, 5–15.
57. Chang, C., Metzger, C.D., Glover, G.H., Duyn, J.H., Heinze, H.J., and Walter, M. (2013). Association between heart rate variability and fluctuations in resting-state functional connectivity. Neuroimage *68*, 93–104.
58. Barber, A.D., John, M., DeRosse, P., Birnbaum, M.L., Lencz, T., and Malhotra, A.K. (2020). Parasympathetic arousal-related cortical activity is associated with attention during cognitive task performance. Neuroimage *208*, 116469.
59. Candia-Rivera, D., Catrambone, V., Thayer, J.F., Gentili, C., and Valenza, G. (2022). Cardiac sympathetic-vagal activity initiates a functional brain–body response to emotional arousal. Proc. Natl. Acad. Sci. USA *119*, e2119599119.
60. Leopold, D.A., Murayama, Y., and Logothetis, N.K. (2003). Very Slow Activity Fluctuations in Monkey Visual Cortex: Implications for Functional Brain Imaging. Cerebr. Cortex *13*, 422–433.
61. Golanov, E.V., Yamamoto, S., and Reis, D.J. (1994). Spontaneous waves of cerebral blood flow associated with a pattern of electrocortical activity. Am. J. Physiol. *266*, R204–R214.
62. Golanov, E.V., and Reis, D.J. (1995). Vasodilation evoked from medulla and cerebellum is coupled to bursts of cortical EEG activity in rats. Am. J. Physiol. *268*, R454–R467.
63. Yeung, N., Botvinick, M.M., and Cohen, J.D. (2004). The Neural Basis of Error Detection: Conflict Monitoring and the Error-Related Negativity. Psychol. Rev. *111*, 931–959.
64. van Veen, V., and Carter, C.S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. Physiol. Behav. *77*, 477–482.
65. Dmochowski, J.P., Bezdek, M.A., Abelson, B.P., Johnson, J.S., Schumacher, E.H., and Parra, L.C. (2014). Audience preferences are predicted by temporal reliability of neural processing. Nat. Commun. *5*, 4567.
66. Meshulam, M., Hasenfratz, L., Hillman, H., Liu, Y.F., Nguyen, M., Norman, K.A., and Hasson, U. (2021). Neural alignment predicts learning outcomes in students taking an introduction to computer science course. Nat. Commun. *12*, 1922.
67. Cohen, S.S., Madsen, J., Touchan, G., Robles, D., Lima, S.F.A., Henin, S., and Parra, L.C. (2018). Neural engagement with online educational videos predicts learning performance for individual students. Neurobiol. Learn. Mem. *155*, 60–64.
68. Critchley, H.D., Mathias, C.J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B.K., Cipolotti, L., Shallice, T., and Dolan, R.J. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. Brain *126*, 2139–2152.
69. Joshi, S., and Gold, J.I. (2022). Context-dependent relationships between locus coeruleus firing patterns and coordinated neural activity in the anterior cingulate cortex. Elife *11*, e63490.
70. O'Sullivan, J.A., Power, A.J., Mesgarani, N., Rajaram, S., Foxe, J.J., Shinn-Cunningham, B.G., Slaney, M., Shamma, S.A., and Lalor, E.C. (2015). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. Cerebr. Cortex *25*, 1697–1706.
71. Gehmacher, Q., Schubert, J., Schmidt, F., Hartmann, T., Reisinger, P., Rösch, S., Schwarz, K., Popov, T., Chait, M., and Weisz, N. (2023). Eye movements track prioritized auditory features in selective attention to natural speech, . Preprint at BioRxiv. https://doi.org/10.1101/2023.01.23.525171.
72. Jin, P., Zou, J., Zhou, T., and Ding, N. (2018). Eye activity tracks task-relevant structures during speech and auditory sequence perception. Nat. Commun. *9*, 5374.
73. Galvez-Pol, A., McConnell, R., and Kilner, J.M. (2020). Active sampling in visual search is coupled to the cardiac cycle. Cognition *196*, 104149.
74. Elofsson, U.O.E., von Schèele, B., Theorell, T., and Söndergaard, H.P. (2008). Physiological correlates of eye movement desensitization and reprocessing. J. Anxiety Disord. *22*, 622–634.
75. Wilson, D.L., Silver, S.M., Covi, W.G., and Foster, S. (1996). Eye movement desensitization and reprocessing: Effectiveness and autonomic correlates. J. Behav. Ther. Exp. Psychiatr. *27*, 219–229.
76. Popov, T., and Staudigl, T. (2023). Cortico-ocular coupling in the service of episodic memory formation. Neurobiology *227*, 102476. https://doi.org/10.1101/2022.11.09.515902.
77. Kelly, S.P., Lalor, E.C., Reilly, R.B., and Foxe, J.J. (2006). Increases in Alpha Oscillatory Power Reflect an Active Retinotopic Mechanism for Distracter Suppression During Sustained Visuospatial Attention. J. Neurophysiol. *95*, 3844–3851.
78. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. Nat. Neurosci. *22*, 1677–1686.
79. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science *364*, 255.
80. Schloegl, A., Ziehe, A., and Müller, K.R. (2009). Automated ocular artifact removal: comparing regression and component-based methods. Nat. Prec., 1. https://doi.org/10.1038/npre.2009.3446.1.
81. Ki, J.J., Parra, L.C., and Dmochowski, J.P. (2020). Visually evoked responses are enhanced when engaging in a video game. Eur. J. Neurosci. *52*, 4695–4708.
82. Iotzov, I., and Parra, L.C. (2019). EEG can predict speech intelligibility. J. Neural. Eng. *16*, 036008.
83. Birbaumer, N., Elbert, T., Canavan, A.G., and Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. Physiol. Rev. *70*, 1–41.
84. Nikulin, V.V., Fedele, T., Mehnert, J., Lipp, A., Noack, C., Steinbrink, J., and Curio, G. (2014). Monochromatic Ultra-Slow (~0.1Hz) Oscillations in the human electroencephalogram and their relation to hemodynamics. Neuroimage *97*, 71–80.
85. Lin, Z., Chen, M., and Ma, Y. (2013). The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. J. Struct. Biol. *181*, 116–127.
86. Parra, L.C., Haufe, S., and Dmochowski, J.P. (2019). Correlated Components Analysis - Extracting Reliable Dimensions in Multivariate Data. Neuron. Behav. Data Anal. Theory *2*, 7125.
87. Cohen, M.X. (2014). Analyzing Neural Time Series Data: Theory and Practice (MIT Press).
88. Geweke, J.F. (1984). Measures of Conditional Linear Dependence and Feedback between Time Series. J. Am. Stat. Assoc. *79*, 907–915.
89. SUBTLEXus. Universiteit Gent https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed data	This paper	10.17605/OSF.IO/3SW5N
Software and algorithms		
Processing and analysis scripts	This paper	10.17605/OSF.IO/3SW5N

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Jens Madsen (jmadsen@ccny.cuny.edu).

Materials availability

This study did not generate unique reagents.

Data and code availability

- Access to data to produce figures in this manuscript can be found in https://osf.io/3sw5n/.
- Access to code that produce each of the figures in this paper see https://osf.io/3sw5n/.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All experiments were carried out at the City College of New York with the approval from the Institutional Review Boards of the City University of New York. Documented informed consent was obtained from all participants at the start of the experiment. Experiment 1 had N = 32 participants (16 Female, age 19–36, M = 23.69, SD = 4.42; 3 subjects were removed due to bad timing, missing or bad signal quality). Experiment 2 had N = 43 subjects participating (31 females, age 18–30, M = 21.58, SD 2.81 years; 5 subjects were removed due to bad timing, missing or bad signal quality). These sample sizes are comparable to our previous studies showing effects of audiovisual and auditory narratives on EEG, heart rate, pupil and gaze position.[33,49]

METHOD DETAILS

Auditory stimuli

Auditory narratives

The stories, which we used in a previous study[37] were autobiographical narratives selected from the StoryCorps project aired on National Public Radio (NPR) and the New York Times Modern Love series (N = 10, 2–5 min each, total duration of 27 min). During the presentation of the audio a gray background was presented on the screen for Experiment 1 and an isoluminant fixation cross was added at the center of the screen for Experiment 2. The stories were selected to be engaging and have an emotional narrative. 5 stories came from the New York Times' Modern Love episodes: ''Broken Heart Doctor'', ''Don't Let it Snow'', ''Falling in Love at 71'', ''Lost and Found'', and ''The Matchmaker'', and 5 stories came from StoryCorps' animated shorts: ''Eyes on the Stars'', ''John and Joe'', ''Marking the Distance'', ''Sundays at Rocco's'' (depicted in Figure 1A), and ''To R.P. Salazar with Love'').

Scrambled audio

To test the effect of low-level sound fluctuations we also generated a ''Scrambled'' version of the narratives (14 min total). This noise-like audio signal preserves the sound envelope and overall spectrum of the speech, but removes the temporal modulation of the spectrum and randomizes the phase to remove harmonics of voiced speech.

Procedure

Participants listened to the stories seated comfortably in a sound-attenuated booth with white fabric walls and normal ambient LED lighting placed around the subject. The stories were played through stereo speakers placed at 60° angles from the subject next to the 27″ monitor, all at a distance of approximately 60 cm from the subject. The procedure of the experiment can be seen on Figure 1.

Attentive and distracted listening

In the attentive condition participants were asked to attend to the story while looking at the screen (Experiment 1) or keep fixation on the isoluminant fixation cross centered on the screen (Experiment 2). In the distracted condition participants listened to the stories again but they were asked to count backwards silently in their mind in steps of 7, starting from a random prime number between 800 and 1000. This condition ensures that the stimuli and surroundings are identical, only changing the participants cognitive state. In Experiment 2 the ''Scrambled'' audio was presented immediately after the ''Attentive'' condition and subjects asked to listen to it as in the ''Attentive'' condition.

Controlled breathing

To test the bottom-up effect of heart rate, we asked participants to breathe synchronously at a rate of 0.1Hz by showing them a waveform of the desired breathing pattern while looking at a centered fixation cross. To eliminate any visual input while they carried out the synchronized breathing task they were asked to continue breathing at the same rate while a uniform gray screen was presented. Data was analyzed during 60 s of this free breathing period (Experiment 1). In the second experiment the free breathing period was increased to 90 s and subjects were asked to look at a centered fixation cross (equal luminance to the gray background).

Controlled luminance

To test the bottom-up effect of pupil size, we asked participants to keep their gaze centered on the screen while we displayed a slowly flashing screen going from white to black (in 5s and back to white in another 5s, so a frequency of 0.1Hz) for a total of 60 s. Prior to the start of the flashing screen we displayed a uniformly gray screen for 60 s. To remove any transient effect of the pupillary response we did not analyze the initial 5 s of the condition.

Controlled saccades

To test the bottom-up effect of eye movements, participants were asked to direct their gaze to a sequence of dots shown on a screen one at a time in 4 different conditions. In each condition dots were either shown for 0.5, 1, 2 or 5 s aiming to produce a saccade rate of 2, 1, 0.5 or 0.2Hz respectively. For each condition a sequence of 9 dots was used (repeated multiple times), which was the same sequence across participants of randomly chosen dot positions, sampled from a 9-point grid (3 by 3 dot design with equidistant points on the screen). To ensure each condition lasted around 90 s the 4 different conditions (0.5, 1, 2 or 5 s) the sequence of 9 points was repeated 20, 10, 5 and 4 times respectively. To minimize any pupillary response due to luminance change when participants were saccading between dots, each dot was chosen to be isoluminant (equal luminance) to the gray background screen.

Questionnaire

After subjects listened to all stories in the attentive condition and carried out control conditions they were asked to answer questions about each story (8–12 questions per narrative, total 72 questions). The list of questions is available in a previous study.[37]

QUANTIFICATION AND STATISTICAL ANALYSIS

Alignment of multimodal recordings

For segmentation of the physiological signals we used common onset and offset triggers, in addition, a beep sound was embedded right before and after each story which were recorded using a StimTracker (Cedrus) to ensure precise alignment across all subjects. To enable all modalities to be on the same timescale we used the lab streaming layer (LSL) protocol. In addition triggers were sent to both the eye tracking and ExG recording systems, by collecting the timestamps from each system when the triggers were sent, we can estimate a linear regression model converting timing of triggers from one modality to the other. This is specifically important when we use the EyeLink system as the quality of the data in ''Streaming'' mode could be diminished.

Recording and preprocessing of gaze position, gaze variation and pupil size

Gaze position and pupil size were recorded using the Eyelink 1000 eye tracker (SR Research Ltd. Ottawa, Canada) with a 35mm lens at a sampling frequency of 500 Hz. Subjects were instructed to sit still while the experiment was carried out, but were free to move their heads, to ensure comfort (no chin rest was used). A standard 9-point calibration scheme was used utilizing manual verification. Stable pupillary responses were ensured by adjusting the background color, dots used for calibration or interventions, or text in the instructions to have equal luminance (iso-luminant). Blinks and saccades were detected using the algorithm of the eye tracker. Artifacts, blinks and 100ms before and after were filled with linearly interpolated values. Effects of blinks and saccades on the pupil signal were removed using linear regression. To regress out the effects of the saccades and blinks on the pupil signal we first downsampled the pupil signal to 100 Hz, the Temporal Response Function was then estimated for saccade and blink offsets in the interval from -4 to 5 s independently (this is the region we have found the effects of saccades and blinks on the pupil size. The linear effects of these two TRFs were then regressed out simultaneously as blinks and saccades have interactions. We compute the gaze variation as follows $gaze_{variation} = \sqrt{H(gaze_{horizontal})^2 + H(gaze_{vertical})^2}$, where H indicates the hilbert transform and *gaze* is the gaze position after band-pass filtering. Both the gaze position, gaze variation and pupil size were downsampled to 128 Hz.

Recording and preprocessing of EEG

The EEG was recorded at a sampling frequency of 2048 Hz using a BioSemi Active Two system. Participants were fitted with a standard, 64-electrode cap with electrode placements following the international 10/10 system, where the ground electrodes were

located next to POz. To later remove any facial or eye movements artifacts from the EEG signal, we additionally recorded the electrooculogram (EOG). We used four auxiliary electrodes (one located above and below the right eye, one located to the right of the left eye and right of the right eye). Robust PCA[85] was used to remove artifacts and outliers from both EEG and EOG data, and subsequently both signals were digitally band-pass filtered (0.3 Hz–64 Hz cutoff) and down-sampled to 128 Hz. Bad EEG electrode channels were identified manually and replaced with interpolated channels. The interpolation was performed using the 3D cartesian coordinates from the electrode cap projected onto a plane using all surrounding "good" electrodes. Any faulty EOG channels were removed, e.g., if no signal was recorded or the power exceeded 3 standard deviations of the other channels. The EOG channels were then used to remove eye-movement artifacts by linearly regressing them from the EEG channels, i.e., least-squares noise cancellation. Effects of heart beats on the EEG signal were removed using linear regression. To regress the effects of heart beats, we first detected R-peaks in the ECG signal (this avoids contaminating the regression with movement artifacts in the ECG reference signals). Then for each participant and electrode an impulse response (TRF) was estimated with latency from -1 to 3 s using impulse train at the R-peaks as input and EEG as output. The predicted EEG is then subtracted. The advantage of using this noise canceling technique over a component subtraction technique, say ICA, is that we are capturing any delayed response to the heart beat, and not just the instantaneous electrical artifact. In each EEG channel, additional outlier samples were identified as values exceeding 4 times the distance between the 25th and the 75th quartile of the median-centered signal, and samples 40 ms before and after such outliers were replaced with interpolated samples using neighboring electrodes.

Recording and preprocessing of ECG

The ECG signal was recorded using a flat electrode placed below the left lumbar region (Left Leg/LL position) with a BioSemi Active Two system at a sampling frequency of 2048 Hz. The ECG signal was detrended using a high-pass filter (0.5 Hz cutoff) and subsequently notch filtered at 60 Hz to remove line noise. Peaks in the ECG corresponding to the R-waves were found using *findpeaks* (built-in MATLAB function). The instantaneous HR is computed for each beat as the inverse of time intervals between subsequent R-wave peaks and interpolated using cubic spline interpolation. To ensure the same sampling frequency for all subjects this instantaneous HR signal is resampled at a regular sampling rate of 128Hz.

Recording and preprocessing of respiration and respiratory volume

The respiration signal was recorded at a sampling frequency of 2048 Hz using the BioSemi Active Two system by a SleepSense 1387 Respiratory Effort Sensors, which captures the tension on a belt worn around the chest of the subject. The polarity of the signal was detected using peaks in the respiration signal (deep breaths) and inverted where necessary to ensure the correct phase of the signal. To compute the respiratory volume signal, we compute the envelope of the respiration signal by taking the absolute value of the hilbert transformed signal.

Inter-subject correlation (ISC)

To determine whether a given factor drives responses we measured the temporal correlation of these signals between pairs of subjects. For most signals this is simply the Pearson's correlation coefficient. For the scalp potentials (EEG) we use correlated-component analysis.[86]

We can compute the inter-subject correlation for the gaze position, gaze variation, pupil size, heart rate, saccade rate, respiration, respiratory volume signals by following 3 steps: 1. computing the Pearson's correlation coefficient between a single subject's signal (each of the modalities independently) and that of all other subjects while they listened to a story or carried out the control experiments. 2. a single ISC value for that subject was obtained by averaging the correlation values between that subject and all other subjects. 3. The two first steps are then repeated for all subjects, resulting in a single ISC value for each subject. For ISC of gaze position we compute the ISC in the horizontal and vertical gaze direction using the procedure as described above separately. To obtain one single ISC value we average the ISC for the horizontal and vertical directions.

The advantage of this technique is that we are not making any assumptions about which feature of the narrative drives these signals, but rather, only assume that the features we captured affect these signals similarly in different subjects. As such, we can capture potentially complex non-linear responses of the signals in questions to the various drivers, despite "only" measuring correlation. In contrast, the WSC and TRF approaches, which also measure correlation, are limited to first-order (linear) relationships between the two signals that are analyzed.

Frequency resolved inter-subject correlation (synchronization spectrum)

We performed a frequency analysis to investigate at which timescale the recorded signals correlated between subjects. Each signal from each subject was band-pass filtered using 5th order Butterworth filters with logarithmic spaced center frequencies with a bandwidth of 0.2 of the center frequency. The ISC was computed for each subject in each frequency band for all narratives. To obtain a single ISC value per frequency band we average ISC values for all narratives and subjects.

Frequency resolved within-subject correlation (WSC)

To capture the correlation between different signal modalities we measure the correlation between signals in different frequency bands (coherence). This is done within subjects, overcoming the limitation of ISC, which requires a common driver to synchronize

signals between subjects. Thus, WSC can be measured in the absence of a common driving force such as during rest. But also, note that if there is a common driving force, WSC may measure a significant correlation that is due to this common factor, and not indicative of a direct causal link between the signals. To measure WSC signals are band-pass filtered (using the same filters as used for the frequency resolved ISC) and correlation is measured with real and imaginary components of the Hilbert transform capturing any potential time delays between signals.

We can do this procedure between one dimensional signals, but for neural signals we follow the procedure as in.[34] We utilize least-squares regression between the EEG signals predicting each of the peripheral signals used in this study. These linear projections of this sort are conventionally referred to as components of the EEG. WSC is then Pearson's correlation between the predicted signal (from the EEG signal) and the peripheral signal in question. In the cases where we differentiate between conditions, we compute component weights and WSC in each condition separately.

WSC in all instances was computed on test data using leave-one-subject out cross-validation (i.e., regression parameters are estimated including all subjects except for the test subject; WSC is then computed on the test subject; this train-test process is then repeated for all subjects.) Statistical significance is established in the same way as for ISC as described in previous sections.

Cluster statistics for contrast between conditions in frequency resolved correlation

To determine significant differences in spectra in different conditions e.g., between conditions (e.g., Figures 2D and 3C), or control conditions and resting state (e.g., Figures 5D and 5E) we use shuffle statistics as follows. Since different frequency bands in the frequency-resolved ISC or WSC are not independent, we use one-dimensional cluster statistics including random shuffles to correct for multiple comparisons following an established procedure.[87] The procedure involves four steps: 1) take the difference between the spectra in two different conditions (e.g., attentive and the distracted) for each subject. 2) compute a one-sample t test on the difference for each frequency band. 3) clusters are identified as consecutive frequency bands with p values below 0.01. The t-values within each cluster are then summed. 4) run 10,000 permutations in which we randomly change one-half of the subjects sign of difference between the spectra computed in step 1. Steps 2–3 are then repeated while keeping the sum of t-values of the largest cluster. Finally, we compare the clusters' t-values obtained in step 3 with the distribution of permuted cluster t-values obtained in step 4. Clusters with larger than 99% (corresponding to p value<0.01) of the permuted distribution were considered significant after multiple comparison cluster correction. Note that for the contrast in attention using the EEG data, all data was used for optimization (of components of the EEG) including attentive and distracted conditions, so that any difference between conditions can not be due to the optimization procedure. In the results section we report the number of subject N, number of frequency bands tested k, and frequency range of the cluster(s) that are significant at $p < 0.01$. This frequency range is also shown in the figures as light gray shading for differences between conditions, or bold lines for differences in ISC from zero.

Cluster statistics for significance of frequency resolved correlation

To determine if frequency-resolved ISC, SRC and WSC values are significantly different from zero we use a similar cluster statistic as above. For cases involving EEG (correlated component analysis and regression) we use test data to avoid upwards bias due to optimization, which is performed on separate training data. The shuffling and cluster correction procedure consists of the same 4 steps as above, except that we divide subjects in two equal size groups at random. The premise of this is that values around zero will not differ significantly if placed at random in two different groups.[87] Reporting of significant clusters in the results section is the same as above.

Temporal response function (TRF)

To capture the specific time delay of correlation between signals we use the classic least-squares linear systems-estimation of the impulse response, which convolves from one signal to predict the another. We allow for acausal delays to determine which signal lead and which signal follows, and estimate the filters in both directions (say $x(t) = h_{xy}(t) * y(t)$ and $y(t) = h_{yx}(t) * x(t)$, where * represents a convolution. Non-zero coefficient in positive lags of $h_{yx}(t)$ indicates that fluctuations in x(t) lead fluctuations in y(t), and vise versa, non-zero coefficient in positive lags of $h_{xy}(t)$ indicates that y(t) leads x(t). This is the basis for Granger ''causality''.[88] However, Granger causality is confounded if there are unobserved common factors. Since that is typically the case here, we refrain from using Granger causality as a tool, and simply report which signal is leading or following another.

Significance of temporal response function

To determine which time delays of correlation between signals are significant for the temporal response function we use shuffle statistics. 1) We compute the filter between the input and output signal (as described above). 2) We then repeat the procedure but now using a circular time-shifted version of the input signal. The delay is chosen randomly from a distribution with a maximum delay of the duration of the input signal. We repeat procedure 2 using 10.000 random time-shifts allowing us to discern p values down to 0.001. To determine whether a specific time delay is significant we estimate the p value for each time delay by comparing the filter estimated in step 1 to the surrogate distribution of filters found in step 2. I.e. how many of the surrogates had a value below the ones found in step 1. We correct the p values found for each time delay using False Discovery Rate (FDR). We use a p value of 0.01 to determine whether a specific time delay is significant. We report significant time delays as above, except the figures, significant delays are indicated by making that portion of the TRF a bold line.

Speech features

To investigate which aspects of the speech signal our brain and body track, we construct various features of the speech signal. First we used the sound envelope to capture low level sound fluctuations following previous work on ''speech tracking''.[39] The envelope is computed as the absolute value of the Hilbert transform of the speech signal, and is down-sampled to 16Hz. We also used word onset (or offset) as a higher level feature of the speech stimulus. For regression, this signal is a pulse train sampled at 16 Hz with a unit pulse at the location of word onset or offset. Finally, we used the bit-rate as a measure of lexical difficulty. We compute for every word the probability of occurrence of that word in the SUBTLEXus lexicon[89] and take -log(2) of that probability and divide that by the inter-word interval. We then subtract the mean of this across all words so that the bit-rate measures if a word has a bit-rate higher or lower than the mean. The corresponding regressor is again a pulse train at the offset of words with pulses scaled by that bit-rate value.