Multiscale neural dynamics for multisensory integration

Daniel Senkowski[1] and Andreas K. Engel[2]

[1] Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Psychiatry and Neurosciences, Charitéplatz 1, 10117 Berlin, Germany

[2] Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany

Corresponding author: Andreas K. Engel, ak.engel@uke.de

Abstract

Carrying out any everyday task, be it driving in traffic, conversing with friends, or playing basketball, requires rapid selection, integration, and segregation of stimuli from different sensory modalities. To date, even the most advanced artificial intelligence-based systems are unable to replicate the multisensory processes that our brain routinely performs. How neural circuits carry out these processes is still not well understood. Here, we discuss recent findings that shed new light on the neural mechanisms underlying multisensory integration (MI). We present a new concept of MI that emphasizes the critical role of neural interactions at multiple timescales, allowing, for example, the simultaneous integration and segregation of information in different time windows. We review recent findings that highlight key mechanisms in the multi-timescale network dynamics mediating MI: power modulations, phase-resetting, phase-amplitude coupling and dynamic functional connectivity. We then discuss scenarios on how multi-timescale processes may coordinate MI in a flexible and adaptive manner. Research on multiscale dynamics may benefit from advances in modeling and neurostimulation techniques and may have implications for multisensory neurotechnologies. The findings discussed here strongly support the notion that multi-timescale interactions constitute a key mechanism in MI.

Introduction

Why do we understand speech better when we watch a person's lip movements? Why does food have little savor if we are unable to process olfactory information[1]? These are questions that are being addressed in research on multisensory integration (MI) which focuses on investigating how information from the different sensory modalities interacts to influence perception and behavior. A key feature of MI is that responses to a multisensory stimulus typically differ from the combined responses of corresponding unisensory stimuli[2,3]. There are also differences in how the brain processes and integrates stimuli within a single sensory modality and across modalities (BOX 1). Studies in the past decades have elucidated a number of factors shaping MI, including bottom-up effects, such as the spatiotemporal congruency of stimuli and the finding of reliable MI effects when the unisensory constituent stimuli elicit weak neural responses[3], the weighting of sensory evidence in relation to the appropriateness for the task at hand[4], and top-down mechanisms such as attention[5]. In this Perspective, we focus on temporal aspects of the mechanisms underlying MI and discuss the concept that MI involves neural dynamics at multiple interacting timescales. Compared to other aspects of MI, the role of neural processes at different timescales involved in MI are much less well understood.

The temporal characteristics of the stimuli that activate peripheral receptor cells, such as sound waves, odor molecules or skin touch, differ considerably between sensory systems[6,7]. Our senses encode intensities and time courses of these stimuli and transduce them into electrical brain signals at timescales that can diverge by orders of magnitude across different systems (FIG. 1a). Furthermore, in all sensory organs, stimuli are encoded concurrently by distinct cell populations at different temporal resolutions. Thus, MI requires comparison and linking of information not only between different sensory systems, but also across different timescales.

At the next processing stages of sensory pathways, the neural dynamics also span a broad variety of timescales (FIG. 1a). Subcortical and cortical regions exhibit synchronized neural activity over a wide range of frequencies, from slow *oscillations* below 1 Hz to high gamma band oscillations at frequencies above 100 Hz[8]. These oscillatory rhythms are involved in many aspects of sensory and cognitive processing[8–15] and it has been hypothesized that they may provide *time windows* of different duration for the chunking or linking of information[10,16–18]. Furthermore, analysis of the *autocorrelation* of neural activity has revealed characteristic timescales of cortical processing[19–23] (FIG. 1b). These *intrinsic neural timescales* seem to differ across processing stages of cortical pathways, with early sensory areas exhibiting shorter timescales and higher-order areas showing progressively longer timescales[19,20,22,23]. Different timescales are also observed in analyses of the dynamic coupling of neural activity across different cell populations[11,17,24–27] (FIG. 1b). Here, there is robust evidence for a plethora of *intrinsic coupling modes* operating on different timescales[26].

Overall, it is well established that the processing of sensory information in the periphery and at central stages involves multiple timescales, i.e., information processing occurs at different temporal resolutions and over a wide range of sensory and neural integration time windows[7,8,19,20,22,23].

Based on these observations, concepts on the functional importance of multi-timescale processing have been advanced that emphasize several distinct, but related aspects. The first aspect concerns the integration of information. Multiple timescales might support the integration of information over time windows of different durations that could be adapted to requirements of the current task[22,23]. Furthermore, it has been suggested that different timescales, reflected in different frequency bands, serve to route different types of information through cortical networks, thus establishing functionally specific links between neural populations[14,28,29]. The second aspect is the segregation of information. It has been proposed that the coexistence of multiple timescales may enhance the capacity of the system by multiplexing, i.e., the parallel processing of complementary information in different channels that could be provided by separate frequency bands[7,30]. A third, closely related aspect is the hierarchical structuring of information. It has been suggested that relations between different timescales may be established by *cross-frequency coupling* of neural oscillations[31,32]. The nesting of multiple processing timescales might lead to the formation of functional hierarchies in cortical networks[19,33]. A fourth aspect concerns the selection of information. Recent concepts postulate the involvement of different neural timescales in prediction and in attentional selection of sensory stimuli[15,34,35]. Both processes require the interplay of top-down and bottom-up information flows which seem to be mediated by different frequency bands[28,36].

These concepts of multi-timescale processing have been substantiated by research on individual sensory modalities (BOX 1). However, multi-timescale concepts have so far not been explicitly considered to account for the dynamic integration and interaction between different sensory systems. In this Perspective, we first review several neural mechanisms that are relevant for multi-timescale dynamics involved in MI. We then describe how multi-timescale dynamics in ongoing activity can influence MI. Subsequently, we discuss the role of multi-timescale MI during cognitive processing. On the basis of this evidence, we will argue for a novel concept of MI that incorporates multi-timescale mechanisms for integration, selection, hierarchical structuring and selection of information. We will describe scenarios that highlight the multi-timescale dynamics underlying MI and allow the testing of predictions derived from the proposed concept.

Neural mechanisms for multiscale MI

The integrative processing of multisensory information requires neural mechanisms that are capable of rapidly adapting to changes in the sensory environment. These must mediate integration across different subsystems of the brain, as well as integration across different timescales. To mediate the linking or segregation of information, the routing of signals and the selection of information, such mechanisms must be able to operate flexibly, both within and between brain areas.

Several neural mechanisms meet these requirements and are likely to be important for multiscale MI (FIG. 1c). These mechanisms include the *amplitude modulation* of neural oscillations in different frequency bands[37–39]; *phase resetting* and *entrainment* of neural signals[40,41]; and dynamic functional coupling of neuronal populations, including *phase coupling*, *envelope coupling*, as well as *phase-amplitude coupling* across frequency bands[26,39,42,43]. These mechanisms are also known to play key roles in multiscale unisensory processing[11,17,26,35], but the neural computations that they perform can differ between unisensory and multisensory processing (BOX 1). Next, we discuss evidence for an involvement of each of these mechanisms in multisensory processing, and we highlight their relevance for multi-timescale MI.

Amplitude modulation

Changes in the amplitude or *power* of neural oscillations (FIG. 1c, top left) during MI have been observed in a large number of studies. For example, studies using multisensory stimuli have shown effects on gamma power (>30 Hz) in sensory cortices[44–48]. This suggests that high-frequency oscillations may provide a neural mechanism with a short time window that supports rapid feed-forward multisensory processing. Further studies using multisensory speech stimuli have shown enhanced frontal theta (4-7 Hz) oscillations in response to semantically incongruent vs. congruent audiovisual stimuli[49] and enhanced frontal delta (<3 Hz) oscillations in response to temporal asynchrony vs. synchrony between visual and auditory speech[50]. This suggests that higher-level multisensory processes may be associated with power modulations in lower frequency bands with longer time windows. Therefore, it is tempting to propose that MI-related power modulations are reflected in a gradient from fast to slow oscillations as one moves from lower to higher level areas. Notably, a similar frequency gradient has been found in human resting state magnetoencephalography data[51] and there is evidence from electrophysiological[52,53] and functional neuroimaging studies[54,55] for a posterior-to-anterior hierarchical organization of multisensory processes. Furthermore, based on the analysis of the autocorrelation of neural signals, it has been suggested that higher levels in sensory processing hierarchies are associated with progressively longer time windows for the integration of information[23,56].

However, several lines of evidence argue against a strictly frequency-dependent hierarchical organization of multisensory processing. First, power modulations of high-frequency oscillations during MI have been found not only in sensory regions but also in higher-order association areas[57–59]. Second, after multisensory learning widespread changes in beta (13-30 Hz) oscillations have been found throughout the cortex, including higher level cortical regions[60,61]. Third, multisensory interactions in low frequency bands have been observed in low level sensory areas[48,62]. Fourth, *crossmodal* prediction errors for incongruent visual and auditory speech stimuli are reflected not only in frontal brain regions[49,50], but also in gamma oscillations in sensory areas[63], contradicting the notion that high-frequency oscillations in sensory regions serve solely rapid feedforward multisensory processes.

In the context of MI, amplitude and power modulations of neural oscillations have been observed across all frequency bands and in all brain subsystems[41,64]. This suggests the presence of multiple frequency-dependent timescales that can operate in parallel to optimize the selection, integration and segregation of sensory information.

Phase resetting

Several studies have shown that the presentation of a stimulus in one sensory modality can crossmodally phase reset (FIG. 1c, bottom left) neural oscillations in sensory cortices of other functional systems[41]. For instance, the presentation of a tactile stimulus leads to a broadband crossmodal phase reset of neural oscillations in auditory cortex with a short latency of less than 100 ms, indicative of feedforward processing[40]. Similarly, the presentation of an auditory stimulus can lead to a rapid phase reset in visual cortex[65] and vice versa[44,66–68]. Crossmodal phase resetting mechanisms in sensory cortices have been shown to be behaviorally relevant[67–69] and they are also influenced by attention[44,70]. Moreover, crossmodal phase resetting has been found to rely on both cortico-cortical projections as well as subcortical feedforward processing[71]. Finally, crossmodal phase resetting has been reported across the entire frequency range of neural oscillations, from long-lasting slow-wave delta oscillations[69,72] to transient high-frequency gamma oscillations[40,44,71].

Thus, the presentation of a stimulus in one sensory modality can induce feedforward phase resetting of neural oscillations in other sensory modalities. Crossmodal phase resetting is found across different frequency bands and with different durations, suggesting that it may serve multiscale information processing to facilitate neural responses and behavior. Feedforward crossmodal phase resetting also appears to be modulated by top-down control[70].

Entrainment

Further research has investigated the entrainment of neural oscillations (FIG. 1c, middle) and its effect on crossmodal processing[73–77]. One study showed that auditory rhythmic delta stimulation leads to an entrainment of slow wave oscillations in visual cortex and a concomitant phase-dependent modulation of visual target sensitivity[76]. In another study, two auditory streams of artificial vowels, each with different pitch and timbre and independently low-pass amplitude modulated, were presented together with a continuous luminance-modulated visual stimulus that matched the amplitude of one of the two streams[75]. Presentation of the visual stimulus enhanced local oscillations in the auditory cortex specifically for the temporally coherent but not for the non-coherent auditory stream. Similarly, a task-irrelevant visual stimulus stream can crossmodally facilitate target detection in a temporally coherent auditory stream compared to non-coherent auditory stream[78]. Furthermore, for crossmodal audio-visual influences, temporally coherent audiovisual streams have been found to facilitate visual target detection, whereas temporally desynchronized audiovisual streams have been found to impair visual target detection under specific conditions[79]. Recently, a study presented continuous 4-Hz modulated tactile and auditory stimuli and showed that in-phase multisensory stimulation enhanced auditory steady-state responses, whereas anti-phase stimulation reduced these responses compared to auditory stimulation alone[80].

The available data suggest that crossmodal entrainment at multiple timescales may provide a neural mechanism that can enhance or impair crossmodal processing, depending on the temporal relationship of the multisensory stimulus streams. Crossmodal entrainment also appears to be well suited for top-down attentional modulation during multisensory processing[41].

Functional coupling

Functional connectivity is abundant in brain networks and can involve both the phase and amplitude of neural signals, giving rise to phase coupling, envelope coupling as well as phase-amplitude coupling[25,26,31] (FIG. 1c, middle and right). A number of studies in humans and in animal models have demonstrated changes in functional connectivity at different timescales in relation to multisensory processing. In the human brain, phase coupling in slow (theta, delta) frequency bands has been demonstrated in sensorimotor networks during audio-visual MI[68,74]. Furthermore, phase coupling at alpha (8-12 Hz) and beta frequencies was observed between the involved sensory cortices during the auditory and tactile *double-flash illusion*[43] and between temporal and fronto-parietal cortex during the *McGurk effect*[81]. During visuo-tactile pattern matching, task related phase coupling has been observed in sensorimotor networks in the theta, alpha, and beta bands[39]. In the same study, long-range alpha-beta phase-amplitude coupling in sensorimotor areas was also found to be predictive of performance. Studies in monkeys have revealed phase coupling in the beta and gamma band between auditory and higher-order temporal cortex during audio-visual MI[82–84]. In rats, visual-tactile sensory processing has been shown to lead to enhanced phase coupling between visual and somatosensory cortex in the theta and beta band[85]. Another study demonstrated that auditory-olfactory associations are accompanied by theta and delta band coupling between auditory and piriform cortex[86]. In ferrets, visual-auditory processing has been shown to be associated with both phase and envelope coupling in multiple frequency bands in networks comprising visual, auditory and parietal areas[42]. While most of the studies demonstrate effects related to phase coupling, there is less evidence for a role of envelope coupling in MI[42].

Taken together, these data show a relation of coupling in multiple frequency ranges to multisensory tasks, supporting the notion that MI requires multi-timescale dynamics not only in terms of local processing, but also with regard to long-range network interactions. Given the distributed nature of cortical sensory systems, such interactions are highly suited to dynamically link information processed in different modalities. Much of the coupling dynamics observed during MI seem intrinsically generated. In the case of rhythmic sensory stimulation, entrainment can lead to phase coupling of the neural populations involved in stimulus processing[74]. However, in most of the studies discussed here the coupling did not result from entrainment by external stimuli but, rather, emerged from intrinsic mechanisms in the respective networks[26].

To further understand how amplitude modulation, phase resetting, entrainment and functional coupling could support multiscale MI, we next discuss how intrinsic network dynamics can shape MI at multiple timescales.

Ongoing multiscale dynamics

Do ongoing intrinsic network dynamics shape MI? Recent studies suggest that ongoing neural dynamics, occurring at different timescales, may indeed influence multisensory processing (FIG. 2). An obvious way to investigate this is to ask how prestimulus activity, as reflected in neural oscillations, affects MI[39,53,81,87–93]. Research on prestimulus neural oscillations and multisensory perception has revealed effects in several frequency ranges. Prestimulus theta oscillations have been related to top-down modulation of crossmodal processing. For instance, reduced ongoing mediofrontal theta oscillations have been found to be associated with an enhancement of the ventriloquist illusion[90]. In addition, prestimulus alpha oscillations have been shown to influence the sensitivity to temporal lags between inputs in different modalities and the perception of the temporal sequencing of audiovisual stimuli[87,93]. Furthermore, spontaneous fluctuations in the power of prestimulus beta oscillations have been shown to predict audio-visual integration in the sound-induced flash illusion[89,91] (FIG. 2a) and in the McGurk effect[81]. In both cases, stronger beta power is associated with enhanced illusory percepts, possibly reflecting an increased predisposition of sensory and higher-order areas for MI[81,89,91]. Hence, these results show that fluctuations in the power of prestimulus neural oscillations with different frequencies, reflecting different timescales, influence upcoming multisensory processing. In the following, we discuss further aspects of ongoing multiscale dynamics that shape MI. These include the brain state reflected in ongoing activity, multiscale prestimulus coupling, learning-dependent changes, and scale-free features of intrinsic brain dynamics.

MI and brain state

Of particular interest is whether time-stable characteristics in the ongoing brain state, such as individual oscillation peak frequencies, can influence MI. Several studies have addressed whether the individual alpha peak frequency relates to the temporal window for audiovisual integration[94–96]. Cecere and coworkers[95] showed that the individual alpha frequency is positively correlated with the size of the temporal multisensory integration window – a finding that has since been replicated in several studies[96–98]; but see[94]. Hence, similar to what has been proposed for unisensory visual processing[99], it is possible that each alpha cycle in a neural network provides a functional window that allows the integration or segregation of information not only within but also across different sensory modalities[100]. Several studies support this conclusion[53,88,92,93]. For example, the perception of audio-tactile stimuli as temporally separated is accompanied by reduced prestimulus alpha power[92] (FIG. 2b).

Furthermore, a computational modeling study has linked prestimulus alpha power and phase to multisensory perception[53]. The study showed that alpha oscillations are directly related to the observer's *causal prior*, i.e., the belief about the causal structure of multisensory stimuli.

Interestingly, this relationship depended on the stimulus history, suggesting that modulations in the amplitude and phase of alpha oscillations may support flexible adaptation processes linked to assumptions about common or distinct sources of incoming sensory information. Thus, there is strong evidence that alpha oscillations are dynamically related to the width of the multisensory integration window and are associated with the observer's beliefs about the causal structure of the world. The phase and amplitude of alpha oscillations can be dynamically adapted to facilitate the integration or segregation of the stimuli that enter our sensory systems at different timescales. However, there is an ongoing controversy as to whether the prestimulus alpha frequency influences the width of the multisensory integration window[96–98] or not[94].

MI and multiscale prestimulus coupling

There is ample evidence that functional coupling in cortical networks prior to stimulus presentation influences multisensory perception[42,81,91,92,101,102]. Several studies on ambiguous audiovisual perception have shown that intrinsic beta band phase coupling can bias the multisensory percept reported by the observers[81,101]. Along similar lines, the sound-induced flash illusion is enhanced by stronger prestimulus beta band phase coupling between multisensory and unisensory regions[91]. Beta band phase coupling has been suggested to play a role in top-down signaling[14,28] and may provide a mechanism for transmitting predictive signals from multisensory regions to lower-level sensory cortices.

In a study on audiotactile interaction, higher prestimulus phase coupling in the alpha band between prefrontal cortex, intraparietal cortex and sensory cortices was found to predict the integration of auditory and tactile inputs[92] (FIG. 2b). This further supports the notion, discussed above, that alpha oscillations relate to the temporal integration window for the perception of simultaneity across sensory modalities[100]. In an auditory-olfactory matching task, delta band and theta band phase coupling in auditory and olfactory cortices was shown to predict task performance[102]. Furthermore, correct trials were associated with coupling between the phase of slow oscillations in auditory and olfactory cortex and the amplitude of high-frequency oscillations in the olfactory cortex. Recent work in ferrets[42] has provided evidence that the coupling of prestimulus signal phases and amplitudes between visual, auditory and partietal areas influences upcoming multisensory processes. Interestingly, prestimulus phase coupling and amplitude coupling, which reflect interactions at different timescales[26], have been observed to correlate with different features of multisensory responses (FIG. 2c). For example, latencies of multisensory response were more strongly predicted by phase coupling than by prestimulus envelope coupling[42]. Overall, these studies provide clear evidence that prestimulus functional coupling in low and high frequency bands, as well as cross-frequency phase-amplitude coupling within and between sensory and higher cortical areas can influence the processing of upcoming multisensory stimuli.

Learning and ongoing multiscale dynamics

Another important question is how multisensory learning and training affect MI by altering the network structure and, thus, shaping the ongoing network dynamics and the disposition for multisensory stimulus processing. Multisensory training over a few days can have significant effects on multisensory perception, such as narrowing the audiovisual temporal integration window[103]. At the neural level, multisensory training has been shown to change beta and gamma band coupling in a widespread network involving prefrontal, parietal, and visual regions[60].

Learning-dependent changes in MI and multiscale neural dynamics are also found after longer training intervals of months to years, as demonstrated by studies in musicians[104–106]. Compared to non-musicians, musicians have a narrower integration time window for detecting audiovisual asynchronies[105,106] and show a higher susceptibility to multisensory illusions[104]. In addition, musicians show enhanced theta and alpha band coupling between auditory, the sensorimotor, and frontal cortices[107]. Hence, multisensory learning and training significantly shape the multiscale neural network dynamics associated with MI. As a consequence, this influences the way multisensory stimuli are processed and integrated.

Scale-free dynamics

Intrinsic multiscale dynamics have been investigated not only with respect to band-limited processes, such as those reflected in neural oscillations, but also for scale-free neural signals[108–112]. These *scale-free dynamics* can be captured, for example, by analyzing the envelopes of neural signals. These typically exhibit *long-range temporal correlations* and *power-law scaling*, indicating that the ongoing activity fluctuates over a wide range of timescales[111,113]. Such features of scale-free dynamics seem highly interesting, as they are thought to reflect a state of *criticality* in which networks can rapidly adapt to volatile environments[108,111,113] and neural coupling patterns are particularly susceptible to perturbations[112]. Unlike frequency-specific oscillatory signals, scale-free neural dynamics have not been directly linked to MI. Thus, one can only speculate that scale-free dynamics might add flexibility to the multisensory integration process by enhancing the capacity for information processing and increasing the variability of neural coupling[112]. Furthermore, it has been argued that scale-free dynamics can improve phase resetting and entrainment in neuronal networks[112], which could also facilitate the multi-timescale interactions required for MI.

In summary, the studies discussed in this section demonstrate that MI is shaped by features of ongoing network dynamics that are intrinsically generated and reflect the current brain state as well as influences of previous learning. These include prestimulus neural oscillations and dynamic coupling in ongoing activity at multiple timescales. The influence of scale-free neural dynamics on MI remains to be determined. As we will discuss in the next section, multiscale neural dynamics may also play an important role in multisensory cognitive processing. To highlight this, we will consider multiscale dynamics in the context of cognitive functions such as multisensory speech processing, intersensory attention, working memory, and predictive processing.

Multiscale MI during cognitive processing

In real-world situations, multisensory stimuli typically exhibit a complex spatiotemporal structure and carry information on more than one timescale. The selection, integration, and segregation of these stimuli require cognitive functions that are presumably realized by hierarchical processing with different temporal resolutions[7,19,22,23,26,30]. Below we discuss four examples for such cognitive functions that are well investigated in the context of MI. All of these examples demonstrate the operation of neural multiscale mechanisms during cognitive processing of multisensory stimuli.

Multiscale entrainment in speech processing

Tracking of speech information involves a nested hierarchy of brain oscillations[114] and parallel information processing in cortical networks with hierarchically organized timescales[115]. This is particularly evident in the case of audiovisual speech which involves crossmodal predictive mechanisms operating at different timescales[116 117]. Studies investigating the neural mechanisms underlying audiovisual speech processing have shown a crossmodal phase-resetting in auditory cortex that facilitates speech perception[73,116–120]. The crossmodal influence is explained by the fact that lip movements precede auditory speech by about 100-300 ms, i.e., the typical duration of a syllable[121]. A recent study has varied the temporal delays between visual and auditory syllables[122]. Examining the nested hierarchical organization of oscillatory networks, the study found that temporal delays were reflected in modulations of oscillatory dynamics on different timescales (FIG. 3a). Furthermore, the study suggests that global network modulations in slow oscillations can tune nested high-frequency local networks, enabling effective integrative processing of speech information across different timescales. Another important finding was that the phase of crossmodally entrained delta-theta oscillations in a left hemisphere-dominated network adapted to the temporal asynchronies between visual and auditory onsets. This shows that the phase of slow wave oscillations can rapidly adapt to the environmental dynamics, providing a flexible neural mechanism for crossmodal information transfer. Another study[120] used computer-generated audiovisual speech stimuli to induce bistable perception, so that either the visual or the auditory input dominated speech processing. The study found that the phase of theta oscillations in temporal and occipital cortex predicted whether the auditory or the visual input dominated audiovisual speech perception[120].

Overall, these studies clearly demonstrate that multisensory speech processing involves visual-auditory phase-resetting and hierarchically organized cross-frequency coupling within and between cortical networks, enabling integrative information processing with different timescales. It seems likely that these neural mechanisms, which can flexibly adapt to changes in the environment, also play a key role in situations other than speech processing that require the tracking and segmentation of complex multisensory stimuli.

Multiscale dynamics in multisensory attention

There is considerable evidence that multiscale neural dynamics are highly relevant for attentional modulation in multisensory contexts. Several lines of research have highlighted the interplay between attention and multisensory integration[5,123]. In addition, numerous studies have investigated the role of neural oscillations for multisensory attention, reporting a variety of effects within and across different frequency bands[37,44,70,124–128]. For example, spatially[44] and non-spatially[70] attended visual stimuli can induce early crossmodal resetting of neural oscillations in auditory cortex, as found in multisensory target detection paradigms. This shows that top-down attention can facilitate multisensory bottom-up processing. In addition, a salient sound can capture visual attention in a purely bottom-up manner, as evidenced by the desynchronization of occipital alpha oscillations and enhanced performance in a visual perception task[124]. Further studies using multisensory target detection paradigms have investigated neural entrainment effects for rhythmically presented sensory stimuli[37,126,127]. For audiovisual stimuli that were presented at sub-delta frequency, neural entrainment was found at twice the stimulation frequency, which was accompanied by modulations of the alpha amplitude in auditory cortex and cross-frequency coupling between the delta phase and alpha power[126]. Relatedly, in an intersensory visuo-tactile attention paradigm[37], entrainment effects for rhythmic vs. non-rhythmic stimuli on alpha and beta power have been observed in the visual cortex (FIG. 3b). The same study revealed functionally connected network effects of attention in the alpha and beta bands, effects of rhythmicity in delta and beta functional networks, and interaction effects between attention and rhythmicity in the theta band[127].

Taken together, this demonstrates that intersensory attention and rhythmic attentional sampling involve neural network dynamics at multiple timescales. An important mechanism seems to be the top-down modulation of slow oscillations which then facilitate higher-frequency bottom-up processing in sensory areas. The oscillatory mechanisms relevant to multisensory attention, in particular crossmodal phase resetting and cross-frequency coupling, overlap to some extent with those that are also involved in other aspects of MI that we have considered above. As we will discuss next, there are also close relations to multiscale MI mechanisms involved in working memory.

Multisensory working memory

Stimuli entering our sensory system typically appear in a specific context and they are preceded and followed by other related input. Perception relies on the selection, integration, and segregation of information from different sensory modalities in a temporal window encompassing a stimulus. This process requires working memory functions, which are intimately related to attention[129] and multisensory processing[130]. Important neural mechanisms that may link the three functions are theta oscillations and cross-frequency phase-amplitude coupling of theta-gamma oscillations. In brain circuits, the phase, amplitude and functional coupling of theta and gamma oscillations can rapidly adapt to changes in the environment. Theta cycles may provide temporal windows for attentional sampling[131], working memory[132] and MI[39]. In humans, enhanced power and functional coupling in the theta band were found during audiovisual object encoding in a match-to-sample task[133], and theta oscillations are also thought to underlie the influence of memory-load on perception-related multisensory processing[134]. In another study, the rhythmic presentation of task-irrelevant visual shapes at theta frequency produced increased frontoparietal connectivity in the theta band, which was found to improve performance in a subsequent auditory working memory task[77]. Thus, there is ample evidence that theta oscillations provide a critical temporal window for the processing of multisensory information in working memory. In addition to these findings, there is also evidence for a beta band supramodal working memory mechanism in the right prefrontal cortex that serves to maintain stimulus frequency in memory[135].

These studies strongly support the conclusion that the interplay between working memory, attention, and MI allows for the flexible and adaptive processing of multisensory information. The available data point to an important role of theta, beta and gamma oscillations, as well as cross-frequency theta-gamma interactions for the selection, integration and segregation of multisensory stimuli in working memory.

Multiscale interactions in predictive processing

Multiscale dynamics are also evident in multisensory *predictive processing*. Based on previous experience, the brain continuously predicts upcoming events and adapts these predictions through learning in order to reduce prediction errors[36,136]. It has been proposed that the top-down signaling of predictions and the bottom-up propagation of prediction errors involves different timescales[36,112,115,137]. Support for this notion comes from the observation that top-down and bottom-up processes are likely to be mediated by oscillatory signals in different frequency bands[11,27,29,137]. While top-down signaling may primarily involve the alpha and beta band, bottom-up processing of novel information seems to be preferentially mediated by interactions in the gamma band[14,27–29,138,139]. Studies on the neural signatures of predictive processing in MI are consistent with this view. For example, prediction errors in audiovisual speech perception are reflected in increased beta band and gamma band dynamics in multisensory temporal cortex[63]. A relation of beta band coupling to the validity of temporal predictions has also been observed in a visuo-tactile task[127]. Interestingly, lower frequency bands also seem to play a role, especially in tasks involving predictions on the timing of sensory stimuli. Phase resetting of ongoing delta band oscillations was found to be relevant for the temporal prediction in a crossmodal time estimation task[69]. Another study showed that temporal prediction accuracy in a delayed target-detection task is associated with increased power of delta and beta oscillations and that correct temporal predictions are preceded by enhanced delta-beta phase-amplitude coupling[140].

Overall, these studies support the notion that oscillatory signals at multiple timescales play a key role in predictive processing. Further evidence comes from work focusing on other aspects of cortical dynamics. It has been proposed that integration time windows, such as those reflected in the autocorrelation of neural responses, increase along functional cortical hierarchies[19,22,23]. This could generate top-down predictions in different time windows and the updating could be based on prediction errors generated on different timescales[23].

In summary, the evidence reviewed in this section suggests that multiscale dynamics are involved in the cognitive processing of multisensory information. The sampling, processing and prediction of environmental events is mediated by the concurrent operation of mechanisms in different frequency bands. In particular, MI seems to involve phase-amplitude coupling between slow oscillations involved in the tracking or rhythmic sampling of sensory signals and faster oscillations relevant for top-down or bottom-up transmission of information. As discussed next, the available evidence leads us to propose a concept of MI that envisages multi-timescale scenarios for the segregation, integration, crossmodal facilitation, and selection of multisensory information.

A multi-timescale concept for MI

It has long been suggested that temporal patterning of brain signals and dynamic coupling across different neural populations is relevant for multisensory processing[13,40,41,64]. However, with few exceptions[40,137], the concurrent involvement of multiple timescales as well as the possible roles and interrelations of processes at these different timescales have not been explicitly considered. The results reviewed in the preceding sections clearly demonstrate the relevance of multi-timescale dynamics in MI. They show that oscillations in different frequency bands[8,9] as well as functional coupling on multiple timescales[11,17,26] provide key mechanisms for MI.

The empirical findings discussed above demonstrate that neural dynamics in all frequency ranges from delta to the gamma band are involved in MI. For all frequency bands, power modulations, phase resetting, and changes in coupling have been observed in the context of multisensory processing. Relatively few functional aspects seem to differ between frequency bands. For example, entrainment to environmental changes or rhythmic samping of stimuli often seems to involve slow oscillations in the delta and theta range[35,131,141]. Furthermore, alpha- and beta-oscillations seem to be primarily related to top-down processing[14,63], whereas the gamma band seems to be mainly involved in the signaling of novel salient information and in the bottom-up processing of multisensory signals[45,47]. The apparent involvement of the whole spectrum of frequencies in MI invites the suggestion that oscillations at different frequencies mainly provide different time windows for the processing of multisensory information. Importantly, this would apply not only to MI, but also to unisensory processing[18,23,30] (BOX 1).

A critical aspect of multi-timescale dynamics is that processes on different timescales can be functionally linked. As discussed above, oscillations in all frequency bands can be involved in cross-frequency coupling, in particular phase-amplitude coupling[31,112]. For example, coupled delta-beta oscillations are entrained by relevant stimuli and serve to track complex sensory signals[140], and coupled theta-gamma oscillations might serve to rhythmically sample stimuli[35,112]. This suggests the intriguing possibility that cross-frequency coupling may establish hierarchical relationships between different timescales by linking information of higher and lower temporal resolution.

Although a complete picture has not yet emerged, a number of conclusions can be drawn regarding the functional implications of the multi-timescale dynamics that underlie MI. Real-life situations that often consist of competing unisensory and multisensory stimuli can serve to illustrate these implications (FIG. 4). In such situations, bottom-up processing of sensory input consistent with the principles of MI is typically enhanced, and these multisensory stimuli are well shielded from other distracting inputs (FIG. 4a). Top-down attention can facilitate multi-timescale processing of attended stimuli and inhibit processing of competing unattended stimuli (FIG. 4b). Below, we discuss four generalizable scenarios to demonstrate how multi-timescale neural interactions are likely to play a key role in MI, and to illustrate functional implications of multi-timescale dynamics for integration, segregation, crossmodal facilitation, and selection of information during multisensory processing (FIG. 4c-f). In all of these scenarios, the interplay between top-down and bottom-up processing is a key feature that is enabled by multi-timescale dynamics.

Segregation

We propose that multi-timescale dynamics are relevant for the segregation of information across sensory modalities. In the scenario shown in FIG. 4c, two different sensory streams of modality 1 (M1) and modality 2 (M2), resulting in neural response patterns in sensory areas Cx1 and Cx2, respectively, encode information with different temporal resolutions. Top-down processing, mediated by cross-frequency coupling between association cortex (AC) and sensory cortices, is likely to be involved in controlling the information flow[70,137,142]. In the scenario illustrated here, neural oscillations in the association cortex are tuned to facilitate the propagation of information from M2 and to inhibit the flow of information from M1, which allows for effective stimulus segregation. In addition, neural interaction patterns in the prestimulus epoch are also likely to be relevant (FIG. 4c). In this scenario, slow coupling dynamics in prestimulus activity creates a bias for selection of the slower M2 pattern in the stimulus interval[126].

Integration

Multi-timescale dynamics may also serve to integrate sensory signals that differ in their temporal structure on one scale, but are related on another timescale (FIG. 4d). We propose that neural coupling at slower time scales with higher-order association areas may facilitate the integration of sensory information across modalities. In the scenario illustrated in FIG. 4d, this is realized by enhanced functional coupling between the association cortex and regions Cx1 and Cx2, which encode information in different temporal integration windows, but share temporal structures on a slower timescale. In such cases, top-down processing in low-frequency delta to alpha bands may be important for MI, especially when attentional selection processes are involved and when the stimuli are semantically related[37,127] (FIG. 3b). Through cross-frequency coupling (e.g., phase-amplitude coupling), these slower rhythms in higher-order cortex may interact with high frequency oscillations in sensory areas that reflect the bottom-up processing of sensory information[70,126]. Importantly, these low frequency oscillations, which provide long temporal integration windows, seem particularly able to adapt to changes in the environment[69,141]. As illustrated in the schematic in FIG. 4d, the integration of sensory information may be further augmented by coupling of the amplitude envelopes of faster rhythms in sensory regions (Cx1, Cx2).

Crossmodal facilitation

Another relevant multi-timescale scenario can involve crossmodal facilitation by phase resetting (FIG. 4e). A salient stimulus in one sensory modality (M1) can phase reset not only ongoing oscillations in that modality, but also also in another modality (M2)[40,44,65]. For example, in the case of audiovisual language processing, phase resetting from the visual to the auditory system may be augmented by neural oscillations in the association cortex and, thus, optimized so that the auditory input (FIG. 4e, blue boxes in M2) reaches the auditory cortex at an optimal phase. In contrast, auditory input that does not fall in the optimal phase is not amplified (FIG. 4e, magenta boxes). Such a scenario is compatible with available evidence on audiovisual speech integration, which shows that phase resetting does indeed facilitate auditory stimulus processing[118] and that global network modulations in slow oscillations can influence nested high-frequency local networks across different timescales[122] (FIG. 3a). We propose that phase-resetting may constitute an essential component of multi-timescale crossmodal processing.

Selection

A fourth scenario (FIG. 4f) illustrates the relevance of multi-timescale dynamics for multisensory competition and attentional selection. In this scenario, we propose that low frequency oscillations in association cortex are specifically aligned with the temporal integration window of the attentionally selected input in the respective sensory cortices. Through phase-amplitude coupling, this alignment may allow for the efficient integration of multisensory information arriving from different modalities[142,143]. In contrast, the processing of competing irrelevant inputs (FIG. 4f, magenta boxes) would be reduced or at least not amplified, because the respective signals in sensory cortex (Cx2, dashed grey curve) are not optimally aligned with the phase of the slow oscillation in association cortex. In this manner, multi-timescale dynamics may support parallel enhancement and integration of relevant sensory inputs while efficiently suppressing competing irrelevant sensory signals.

Overall, we propose that multi-timescale neural processes have a pivotal role for MI. Multi-timescale processing may be critical for the multiplexed encoding of multisensory information, as well as the integration, segregation, selection and routing of signals. Our view emphasizes the role of intrinsically generated dynamics and that of top-down processing implemented through coupling between slow and faster rhythms. One implication of the concept discussed here is that a more general view on oscillations may emerge which suggests that different frequency ranges do not necessarily mediate strictly distinct functions but, rather, enable flexible processing on different interacting timescales. Depending on the brain subsystem and the context or task in which they are deployed, the mechanisms discussed in this section are likely to serve multiple purposes. For example, the scenario of multi-timescale dynamics underlying selection (FIG. 4f) could mediate multisensory attention in parietal networks, but also context-dependent rule switching in prefrontal circuits.

Outlook

In this Perspective, we have introduced a concept of multisensory processing that emphasizes the critical role of multi-timescale neural dynamics. The findings discussed here are compatible with prominent concepts on multiscale network dynamics and brain functions[11,20,22,23,26,27,112,144]. The multi-timescale view of multisensory communication and integration that we advocate here builds on and goes beyond these frameworks. While our multi-timescale concept of MI is supported by much of the experimental evidence discussed in this article, many of its implications remain to be tested in future studies.

One issue that requires further investigation is the relationship between different timescales. The concept proposed here strongly predicts causal interactions between the mechanisms operating on different timescales. This might enable the emergence of functional hierarchies, e.g., for encoding of part-whole relationships in language processing or in working memory. However, this has not been tested experimentally. Another important unresolved question concerns the relation of scale-specific, i.e., oscillatory, and scale-free features of multi-time-scale dynamics. It has been proposed that the latter can have an important influence on the former[112]. There is evidence from modeling work that critical states of brain networks, as reflected by power-law scaling and long-range temporal correlations, are characterized by intermediate levels of oscillatory coupling, which allow for rapid changes in the network state[112]. However, this remains to be tested experimentally by using appropriate interventions. Furthermore, the potential role of subcortical structures, such as the superior colliculus or the pulvinar, in the emergence of multi-timescale dynamics underlying MI needs to be elucidated. Finally, demonstrating the functional relevance of multi-timescale interactions for MI also implies identifying possible readout mechanisms that use these dynamics to generate behavior. Resolving these questions likely will require the combination of neurophysiological recordings with computational modeling and targeted neurostimulation.

At the behavioral level, the use of modeling approaches, such as the Bayesian framework[145,146] and multisensory correlation detectors[147,148], has significantly advanced our understanding of MI. Only in recent years, attempts have attempts been made to link behavioral models to functional neuroimaging[54,55] and neurophysiological[52,53] data in multisensory paradigms. Such neurally-informed modeling approaches hold great promises for testing some of the open research questions mentioned above, especially when considering relevant details of the microcircuit structure such as the role of neural populations in different cortical layers in generating multi-timescale dynamics[149,150]. Another approach with great potential is whole-brain modeling of virtual brains, which can combine multi-timescale neural dynamics with complex network architectures[151,152]. To date, the whole-brain approach is still limited in accounting for the neurophysiological complexity of neural circuits. Hence, it needs further elaboration before it may be used to study the multiscale dynamics underlying MI. Importantly, computational modeling approaches are also needed that link neural signals across different spatial scales enabling, for example, to infer cellular-level parameters from noninvasive recordings of mass signals[153].

Important advances have been made in brain stimulation and recording techniques that will enhance the study of multiscale neural dynamics in MI. One promising avenue is the use of closed-loop applications using noninvasive brain stimulation techniques such as transcranial alternating current stimulation[154]. It is now possible to target and manipulate brain oscillations in a personalized, phase-dependent manner, e.g., to stimulate precisely at the individual peak or the trough of the oscillatory cycle[155]. To investigate the causal role of multi-timescale dynamics in MI, it may be promising to apply stimulation protocols that combine stimulation in multiple frequency bands, or to use amplitude modulation for alternating current stimulation that mimics the natural dynamics. Animal studies also hold great potential in this regard, as optogenetic stimulation[156,157] can be precisely targeted to brain regions and subpopulations of neurons, thus more specifically addressing the mechanisms underlying multi-timescale neural dynamics.

Investigating the multiscale dynamics underlying MI may also be important for improving our understanding of altered multisensory processing in patients with schizophrenia[158,159] or autism spectrum disorder[160] (ASD). A consistent finding in research on ASD and schizophrenia is an enlarged temporal audiovisual integration window[158,161], and there is evidence for alterations in crossmodal predictive coding in both disorders[49,162]. It is possible that these deficits are related to abnormal multi-timescale neural processing, expressed in aberrant cross-frequency coupling[163] and alterations in scale-free network dynamics[164,165]. However, studies on multiscale neural dynamics underlying alterations in multisensory processing in schizophrenia and ASD remain sparse.

Research on multiscale dynamics and MI likely will also have strong implications for improving multisensory neurotechnologies in sensory impaired individuals and clinical populations. Examples include multisensory stimulation for the treatment of phantom limb pain[166] and neuropathic pain[167], or for the optimization of visuotactile feedback devices for the blind[168], which have already been successfully applied. More recently, there have been advances in the development of brain-computer interfaces, including multisensory stimulation, for the treatment of neuropathic pain[167]. This approach involved the real-time measurement of pain at multiple timescales, as reflected in neural oscillations, using machine learning techniques to control multisensory stimulation. Finally, there is evidence that multiscale neural dynamics should be thoroughly considered in the development and optimization of visuo-tactile feedback devices for congenitally blind individuals[168].

In conclusion, research on multi-timescale neural dynamics for MI has only recently emerged and, as outlined here, many open questions remain to be addressed. The advances in modeling and brain stimulation techniques will help to test predictions that can be derived from the multi-timescale concept for MI we have proposed. Our concept on multi-timescale neural dynamics for MI could have important implications and guide future basic and clinical research in the field of multisensory processing and beyond.

Acknowledgements

D.S. discloses support for this work from the DFG (SE1859/10-1). A.K.E. acknowledges support for this work from the DFG (SFB936-178316478-A2/A3; TRR169-261402652-B1/Z2) and from the European Union (project cICMs, ERC-2022-AdG-101097402). Views and opinions expressed in this article are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

References

1. Auvray, M. & Spence, C. The multisensory perception of flavor. *Conscious. Cogn.* **17**, 1016-1031 (2008).
2. Stein, B. E. & Meredith, M. A. *The Merging of the Senses.* (The MIT Press, 1993).
3. Stein, B. E. & Stanford, T. R. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**, 255-266 (2008).
4. Welch, R. B. & Warren, D. H. Immediate perceptual response to intersensory discrepancy. *Psychol. Bull*. **88**, 638-667 (1980).
5. Talsma, D., Senkowski, D., Soto-Faraco, S. & Woldorff, M. G. The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* **14**, 400-410 (2010).
6. Spence, C. & Squire, S. Multisensory integration: maintaining the perception of synchrony. *Curr. Biol.* **13**, R519-21 (2003).
7. Panzeri, S., Brunel, N., Logothetis, N. K. & Kayser, C. Sensory neural codes using multiplexed temporal scales. *Trends Neurosci*. **33**, 111-120 (2010).
8. Buzsáki, G. & Draguhn, A. Neuronal oscillations in cortical networks. *Science* **304**, 1926-1929 (2004).
9. Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555-586 (1995).
10. Singer, W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* **24**, 49-65, 111 (1999).
11. Engel, A. K., Fries, P. & Singer, W. Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* **2**, 704-716 (2001).
12. Fries, P., Nikolić, D. & Singer, W. The gamma cycle. *Trends Neurosci.* **30**, 309-316 (2007).
13. Senkowski, D., Schneider, T. R., Foxe, J. J. & Engel, A. K. Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* **31**, 401-409 (2008).
14. Engel, A. K. & Fries, P. Beta-band oscillations – signalling the status quo? *Curr. Opin. Neurobiol.* **20**, 156-165 (2010).
15. Jensen, O., Bonnefond, M. & VanRullen, R. An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn. Sci.* **16**, 200-206 (2012).
16. Lisman, J. The theta/gamma discrete phase code occuring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* **15**, 913-922 (2005).
17. Fries, P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**, 474-480 (2005).
18. Jensen, O., Gips, B., Bergmann, T. O. & Bonnefond, M. Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci.* **37**, 357-369 (2014).
19. Hasson, U., Chen, J. & Honey, C. J. Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* **19**, 304-313 (2015).
20. Wang, X.-J. & Kennedy, H. Brain structure and dynamics across scales: in search of rules. *Curr. Opin. Neurobiol* .**37**, 92-98 (2016).
21. Cavanagh, S. E., Hunt, L. T. & Kennerley, S. W. A diversity of intrinsic timescales underlie neural computations. *Front. Neural Circuits* **14**, 615626 (2020).
22. Soltani, A., Murray, J. D., Seo, H. & Lee, D. Timescales of cognition in the brain. *Curr. Opin. Behav. Sci.* **41**, 30-37 (2021).
   This article discusses how information processing on different neural timescales may contribute to sensory and reward integration in the brain.
23. Wolff, A., Berberian, N., Golesorkhi, M., Gomez-Pilar, J., Zilio, F. & Northoff, G. Intrinsic neural timescales: temporal integration and segregation. *Trends Cogn. Sci.* **26**, 159-173 (2022).
   Comprehensive review on the role of intrinsic neural timescales, as measured by the autocorrelation window, for the integration and segregation of sensory input.

24. Singer, W., Engel, A. K., Kreiter, A. K., Munk, M. H., Neuenschwander, S. & Roelfsema, P. R. Neuronal assemblies: necessity, signature and detectability. *Trends Cogn. Sci.* **1**, 252-261 (1997).
25. Siegel, M., Donner, T. H. & Engel, A. K. Spectral fingerprints of large-scale neuronal interactions. *Nat. Rev. Neurosci.* **13**, 121-134 (2012).
26. Engel, A., Gerloff, C., Hilgetag, C. & Nolte, G. Intrinsic coupling modes: multiscale interactions in ongoing brain activity. *Neuron* **80**, 867-886 (2013).
This review summarizes evidence for the functional importance of intrinsically generated neural coupling at multiple timescales.
27. Fries, P. Rhythms for cognition: communication through coherence. *Neuron* **88**, 220-235 (2015).
28. Bastos, A. M. et al. Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* **85**, 390-401 (2015).
29. Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H. & Fries, P. Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron* **89**, 384-397 (2016).
30. Helfrich, R. F. & Knight, R. T. Oscillatory dynamics of prefrontal cognitive control. *Trends Cogn. Sci.* **20**, 916-930 (2016).
31. Canolty, R. T. & Knight, R. T. The functional role of cross-frequency coupling. *Trends Cogn. Sci.* **14**, 506-515 (2010).
32. Hyafil, A., Giraud, A.-L., Fontolan, L. & Gutkin, B. Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends Neurosci.* **38**, 725-740 (2015).
33. Yamashita, Y. & Tani, J. in *Computational and Robotic Models of the Hierarchical Organization of Behavior* (ed Baldassarre, G., M) 47-62 (Springer, 2013).
34. Buschman, T. J. & Kastner, S. From behavior to neural dynamics: an integrated theory of attention. *Neuron* **88**, 127-144 (2015).
35. Fries, P. Rhythmic attentional scanning. *Neuron* **111**, 954-970 (2023).
36. Bastos, A., Usrey, W., Adams, R., Mangun, G., Fries, P. & Friston, K. Canonical microcircuits for predictive coding. *Neuron* **76**, 695-711 (2012).
37. Pomper, U., Keil, J., Foxe, J. J. & Senkowski, D. Intersensory selective attention and temporal orienting operate in parallel and are instantiated in spatially distinct sensory and motor cortices. *Hum. Brain Mapp.* **36**, 3246-3259 (2015).
38. Daume, J., Graetz, S., Gruber, T., Engel, A. K. & Friese, U. Cognitive control during audiovisual working memory engages frontotemporal theta-band interactions. *Sci. Rep.* **7**, (2017).
39. Wang, P., Göschl, F., Friese, U., König, P. & Engel, A. K. Long-range functional coupling predicts performance: oscillatory EEG networks in multisensory processing. *Neuroimage* (2019).
This EEG study demonstrates phase coupling in multiple frequency bands, as well as phase-amplitude coupling, during visuotactile pattern matching.
40. Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A. & Schroeder, C. E. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* **53**, 279-292 (2007).
This study demonstrates broadband crossmodal tactile-auditory phase resetting in the monkey primary auditory cortex.
41. Bauer, A. R., Debener, S. & Nobre, A. C. Synchronisation of neural oscillations and cross-modal influences. *Trends Cogn. Sci.* **24**, 481-495 (2020).
This review highlights the critical role of crossmodal phase resetting and neural entrainment in multisensory processing.
42. Galindo-Leon, E. E., Stitt, I., Pieper, F., Stieglitz, T., Engler, G. & Engel, A. K. Context-specific modulation of intrinsic coupling modes shapes multisensory processing. *Sci. Adv.* **5**, eaar7633 (2019).
This study in the ferret cortex shows that phase and amplitude coupling in different frequency bands predict specific aspects of multisensory processing.

43. Cooke, J., Poch, C., Gillmeister, H., Costantini, M. & Romei, V. Oscillatory properties of functional connections between sensory areas mediate cross-modal illusory perception. *J. Neurosci.* **39**, 5711-5718 (2019).
44. Senkowski, D., Talsma, D., Herrmann, C. S. & Woldorff, M. G. Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Exp. Brain Res.* **166**, 411-426 (2005).
45. Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S. & Woldorff, M. G. Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia* **45**, 561-571 (2007).
46. Quinn, B. T. et al. Intracranial cortical responses during visual-tactile integration in humans. *J. Neurosci.* **34**, 171-181 (2014).
47. Krebber, M., Harwood, J., Spitzer, B., Keil, J. & Senkowski, D. Visuotactile motion congruence enhances gamma-band activity in visual and somatosensory cortices. *Neuroimage* **117**, 160-169 (2015).
48. Ferraro, S. et al. Stereotactic electroencephalography in humans reveals multisensory signal in early visual and auditory cortices. *Cortex* **126**, 253-264 (2020).
49. Roa Romero, Y., Keil, J., Balz, J., Gallinat, J. & Senkowski, D. Reduced frontal theta oscillations indicate altered crossmodal prediction error processing in schizophrenia. *J. Neurophysiol.* **116**, 1396-1407 (2016).
50. Biau, E., Schultz, B. G., Gunter, T. C. & Kotz, S. A. Left motor δ oscillations reflect asynchrony detection in multisensory speech perception. *J. Neurosci.* **42**, 2313-2326 (2022).
51. Mahjoory, K., Schoffelen, J. M., Keitel, A. & Gross, J. The frequency gradient of human resting-state brain oscillations follows cortical hierarchies. *Elife* **9**, e53715 (2020).
52. Cao, Y., Summerfield, C., Park, H., Giordano, B. L. & Kayser, C. Causal inference in the multisensory brain. *Neuron* **102**, 1076-1087.e8 (2019).
53. Rohe, T., Ehlis, A. C. & Noppeney, U. The neural dynamics of hierarchical bayesian causal inference in multisensory perception. *Nat. Commun.* **10**, 1907 (2019).
54. Rohe, T. & Noppeney, U. Cortical hierarchies perform bayesian causal inference in multisensory perception. *PLoS Biol*. **13**, e1002073 (2015).
55. Rohe, T. & Noppeney, U. Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Curr. Biol.* **26**, 509-514 (2016).
56. Murray, J. D. et al. A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661-1663 (2014).
57. Balz, J. et al. Gaba concentration in superior temporal sulcus predicts gamma power and perception in the sound-induced flash illusion. *Neuroimage* **125**, 724-730 (2016).
58. Ozker, M., Schepers, I. M., Magnotti, J. F., Yoshor, D. & Beauchamp, M. S. A double dissociation between anterior and posterior superior temporal gyrus for processing audiovisual speech demonstrated by electrocorticography. *J. Cogn. Neurosci.* **29**, 1044-1060 (2017).
59. Michail, G., Senkowski, D., Holtkamp, M., Wächter, B. & Keil, J. Early beta oscillations in multisensory association areas underlie crossmodal performance enhancement. *Neuroimage* **257**, 119307 (2022).
60. La Rocca, D., Ciuciu, P., Engemann, D. A. & van Wassenhove, V. Emergence of β and γ networks following multisensory training. *Neuroimage* **206**, 116313 (2020).
This MEG study demonstrates changes in stimulus-induced beta- and gamma-band network oscillations after multisensory training.
61. Theves, S., Chan, J. S., Naumer, M. J. & Kaiser, J. Improving audio-visual temporal perception through training enhances beta-band activity. *Neuroimage* **206**, 116312 (2020).
62. Göschl, F., Friese, U., Daume, J., König, P. & Engel, A. K. Oscillatory signatures of crossmodal congruence effects: an EEG investigation employing a visuotactile pattern matching paradigm. *Neuroimage* **116**, 177-186 (2015).
63. Arnal, L. H., Wyart, V. & Giraud, A.-L. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* **14**, 797-801 (2011).

This MEG study shows that crossmodal prediction errors in audiovisual speech processing are expressed in multi-timescale neural dynamics.

64. Keil, J. & Senkowski, D. Neural oscillations orchestrate multisensory processing. *The Neuroscientist* **24**, 609-626 (2018).
This review provides an overview of the functional role of neural oscillations with different timescales for multisensory processing.
65. Mercier, M. R., Foxe, J. J., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H. & Molholm, S. Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. *Neuroimage* **79**, 19-29 (2013).
66. Kayser, C., Petkov, C. I. & Logothetis, N. K. Visual modulation of neurons in auditory cortex. *Cereb. Cortex* **18**, 1560-1574 (2008).
67. Thorne, J. D., De Vos, M., Viola, F. C. & Debener, S. Cross-modal phase reset predicts auditory task performance in humans. *J. Neurosci.* **31**, 3853-3861 (2011).
68. Mercier, M. R., Molholm, S., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H. & Foxe, J. J. Neuro-oscillatory phase alignment drives speeded multisensory response times: an electro-corticographic investigation. *J. Neurosci.* **35**, 8546-8557 (2015).
This human electrocorticography study demonstrates the behavioral relevance of visual-auditory crossmodal phase resetting in auditory cortex.
69. Daume, J., Wang, P., Maye, A., Zhang, D. & Engel, A. K. Non-rhythmic temporal prediction involves phase resets of low-frequency delta oscillations. *Neuroimage* **224**, 117376 (2021).
70. Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C. & Schroeder, C. E. The leading sense: supramodal control of neurophysiological context by attention. *Neuron* **64**, 419-430 (2009).
71. Sieben, K., Roder, B. & Hanganu-Opatz, I. L. Oscillatory entrainment of primary somatosensory cortex encodes visual control of tactile processing. *J. Neurosci.* **33**, 5736-5749 (2013).
This study in rats shows that cortico-cortical interactions contribute to crossmodal phase resetting on multiple timescales in somatosensory cortex.
72. Kuroki, S. et al. Excitatory neuronal hubs configure multisensory integration of slow waves in association cortex. *Cell Rep.* **22**, 2873-2885 (2018).
73. Crosse, M. J., Butler, J. S. & Lalor, E. C. Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. *J. Neurosci.* **35**, 14195-14204 (2015).
74. Clouter, A., Shapiro, K. L. & Hanslmayr, S. Theta phase synchronization is the glue that binds human associative memory. *Curr. Biol.* **27**, 3143-3148.e6 (2017).
75. Atilgan, H. et al. Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron* **97**, 640-655.e4 (2018).
76. Bauer, A. R., van Ede, F., Quinn, A. J. & Nobre, A. C. Rhythmic modulation of visual perception by continuous rhythmic auditory stimulation. *J. Neurosci.* **41**, 7065-7075 (2021).
77. Albouy, P., Martinez-Moreno, Z. E., Hoyer, R. S., Zatorre, R. J. & Baillet, S. Supramodality of neural entrainment: rhythmic visual stimulation causally enhances auditory working memory performance. *Sci. Adv.* **8**, eabj9782 (2022).
78. Maddox, R. K., Atilgan, H., Bizley, J. K. & Lee, A. K. Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners. *Elife* **4**, e04995 (2015).
79. Kösem, A. & van Wassenhove, V. Temporal structure in audiovisual sensory selection. *PLoS One* **7**, e40936 (2012).
80. Fu, X. & Riecke, L. Effects of continuous tactile stimulation on auditory-evoked cortical responses depend on the audio-tactile phase. *Neuroimage* **274**, 120140 (2023).
81. Keil, J., Müller, N., Ihssen, N. & Weisz, N. On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cereb. Cortex* (2012).

82. Ghazanfar, A. A., Chandrasekaran, C. & Logothetis, N. K. Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J. Neurosci.* **28**, 4457-4469 (2008).
83. Maier, J. X., Chandrasekaran, C. & Ghazanfar, A. A. Integration of bimodal looming signals through neuronal coherence in the temporal lobe. *Curr. Biol.* **18**, 963-968 (2008).
84. Kayser, C. & Logothetis, N. K. Directed interactions between auditory and superior temporal cortices and their role in sensory integration. *Front. Integr. Neurosci.* **3**, 7 (2009).
85. Bieler, M., Sieben, K., Cichon, N., Schildt, S., Röder, B. & Hanganu-Opatz, I. L. Rate and temporal coding convey multisensory information in primary sensory cortices. *eNeuro* **4**, ENEURO.0037-17.2017 (2017).
86. Olofsson, J. K., Zhou, G., East, B. S., Zelano, C. & Wilson, D. A. Odor identification in rats: behavioral and electrophysiological evidence of learned olfactory-auditory associations. *eNeuro* **6**, ENEURO.0102-19.2019 (2019).
87. Grabot, L., Kösem, A., Azizi, L. & van Wassenhove, V. Prestimulus alpha oscillations and the temporal sequencing of audiovisual events. *J. Cogn. Neurosci.* **29**, 1566-1582 (2017).
88. Ikumi, N., Torralba, M., Ruzzoli, M. & Soto-Faraco, S. The phase of pre-stimulus brain oscillations correlates with cross-modal synchrony perception. *Eur. J. Neurosci.* **49**, 150-164 (2019).
89. Kaiser, M., Senkowski, D., Busch, N. A., Balz, J. & Keil, J. Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Sci. Rep.* **9**, 5983 (2019).
90. Kaiser, M., Senkowski, D. & Keil, J. Mediofrontal theta-band oscillations reflect top-down influence in the ventriloquist illusion. *Hum. Brain Mapp.* **42**, 452-466 (2021).
91. Keil, J., Müller, N., Hartmann, T. & Weisz, N. Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cereb. Cortex* (2014).
92. Leonardelli, E., Braun, C., Weisz, N., Lithari, C., Occelli, V. & Zampini, M. Prestimulus oscillatory alpha power and connectivity patterns predispose perceptual integration of an audio and a tactile stimulus. *Hum. Brain Mapp.* **36**, 3486-3498 (2015).
93. London, R. E., Benwell, C. S. Y., Cecere, R., Quak, M., Thut, G. & Talsma, D. EEG alpha power predicts the temporal sensitivity of multisensory perception. *Eur. J. Neurosci.* **55**, 3241-3255 (2022).
94. Buergers, S. & Noppeney, U. The role of alpha oscillations in temporal binding within and across the senses. *Nat. Hum. Behav.* **6**, 732-742 (2022).
95. Cecere, R., Rees, G. & Romei, V. Individual differences in alpha frequency drive crossmodal illusory perception. *Curr. Biol.* **25**, 231-235 (2015).
96. Keil, J. & Senkowski, D. Individual alpha frequency relates to the sound-induced flash illusion. *Multisens. Res.* **30**, 565-578 (2017).
97. Bastiaansen, M., Berberyan, H., Stekelenburg, J. J., Schoffelen, J. M. & Vroomen, J. Are alpha oscillations instrumental in multisensory synchrony perception. *Brain Res.* **1734**, 146744 (2020).
98. Venskus, A. & Hughes, G. Individual differences in alpha frequency are associated with the time window of multisensory integration, but not time perception. *Neuropsychologia* **159**, 107919 (2021).
99. VanRullen, R. & Koch, C. Is perception discrete or continuous? *Trends Cogn. Sci.* **7**, 207-213 (2003).
100. Lange, J., Keil, J., Schnitzler, A., van Dijk, H. & Weisz, N. The role of alpha oscillations for illusory perception. *Behav. Brain Res.* **271**, 294-301 (2014).
101. Hipp, J. F., Engel, A. K. & Siegel, M. Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* **69**, 387-396 (2011).
102. Zhou, G. et al. Human olfactory-auditory integration requires phase synchrony between sensory cortices. *Nat. Commun.* **10**, 1168 (2019).
Using human intracranial recordings, this study shows that auditory-olfactory matching involves phase-amplitude coupling between auditory and olfactory cortical regions.

103. Powers, A. R., Hillock, A. R. & Wallace, M. T. Perceptual training narrows the temporal window of multisensory binding. *J. Neurosci.* **29**, 12265-12274 (2009).
104. Bidelman, G. M. Musicians have enhanced audiovisual multisensory binding: Experience-dependent effects in the double-flash illusion. *Exp. Brain Res.* **234**, 3037-3047 (2016).
105. Lee, H. & Noppeney, U. Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proc. Natl. Acad. Sci. USA* **108**, E1441-50 (2011).
106. O'Donohue, M., Lacherez, P. & Yamamoto, N. Musical training refines audiovisual integration but does not influence temporal recalibration. *Sci. Rep.* **12**, 15292 (2022).
107. Klein, C., Liem, F., Hänggi, J., Elmer, S. & Jäncke, L. The "silent" imprint of musical training. *Hum. Brain Mapp.* **37**, 536-546 (2016).
EEG study in musicians showing that sustained training influences multi-timescale network dynamics in the brain.
108. Cocchi, L., Gollo, L. L., Zalesky, A. & Breakspear, M. Criticality in the brain: a synthesis of neurobiology, models and cognition. *Prog. Neurobiol.* **158**, 132-152 (2017).
This review highlights the important role of criticality, which is characterized by activity fluctuations that have no preferred timescale and allow the brain to rapidly adapt to environmental changes.
109. Klar, P., Çatal, Y., Langner, R., Huang, Z. & Northoff, G. Scale-free dynamics of core-periphery topography. *Hum. Brain Mapp.* **44**, 1997-2017 (2023).
110. La Rocca, D., Zilber, N., Abry, P., van Wassenhove, V. & Ciuciu, P. Self-similarity and multifractality in human brain activity: a wavelet-based analysis of scale-free brain dynamics. *J. Neurosci. Meth.* **309**, 175-187 (2018).
111. Palva, J. M., Zhigalov, A., Hirvonen, J., Korhonen, O., Linkenkaer-Hansen, K. & Palva, S. Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proc. Natl. Acad. Sci. USA* **110**, 3585-3590 (2013).
112. Palva, S. & Palva, J. M. Roles of brain criticality and multiscale oscillations in temporal predictions for sensorimotor processing. *Trends Neurosci.* **41**, 729-743 (2018).
This review highlights the importance of scale-free neural dynamics for the role of neural oscillations and cross-frequency coupling in sensorimotor processing.
113. Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M. & Ilmoniemi, R. J. Long-range temporal correlations and scaling behavior in human brain oscillations. *J. Neurosci.* **21**, 1370-1377 (2001).
114. Gross, J. et al. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol.* **11**, e1001752 (2013).
115. Schmitt, L. M., Erb, J., Tune, S., Rysop, A. U., Hartwigsen, G. & Obleser, J. Predicting speech from a cortical hierarchy of event-based time scales. *Sci. Adv.* **7**, eabi6070 (2021).
116. Park, H., Kayser, C., Thut, G. & Gross, J. Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *Elife* **5**, e14521 (2016).
117. Aller, M., Økland, H. S., MacGregor, L. J., Blank, H. & Davis, M. H. Differential auditory and visual phase-locking are observed during audio-visual benefit and silent lip-reading for speech perception. *J. Neurosci.* **42**, 6108-6120 (2022).
118. Mégevand, P. et al. Crossmodal phase reset and evoked responses provide complementary mechanisms for the influence of visual speech in auditory cortex. *J. Neurosci.* **40**, 8530-8542 (2020).
119. Perrodin, C., Kayser, C., Logothetis, N. K. & Petkov, C. I. Natural asynchronies in audiovisual communication signals regulate neuronal multisensory interactions in voice-sensitive cortex. *Proc. Natl. Acad. Sci. USA* **112**, 273-278 (2015).
120. Thézé, R., Giraud, A. L. & Mégevand, P. The phase of cortical oscillations determines the perceptual fate of visual cues in naturalistic audiovisual speech. *Sci. Adv.* **6**, eabc6348 (2020).
121. Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A. & Ghazanfar, A. A. The natural statistics of audiovisual speech. *PLoS Comput. Biol.* **5**, e1000436 (2009).

This study shows that multi-timescale neural responses in the brain, as reflected in different frequency band oscillations, match to the natural features of audiovisual speech.

122. Chalas, N., Omigie, D., Poeppel, D. & van Wassenhove, V. Hierarchically nested networks optimize the analysis of audiovisual speech. *iScience* **26**, 106257 (2023).
This MEG study in humans demonstrates that hierarchically nested oscillatory brain networks track audiovisual speech asynchronies.
123. Noppeney, U. Perceptual inference, learning, and attention in a multisensory world. *Annu. Rev. Neurosci.* **44**, 449-473 (2021).
124. Feng, W., Störmer, V. S., Martinez, A., McDonald, J. J. & Hillyard, S. A. Involuntary orienting of attention to a sound desynchronizes the occipital alpha rhythm and improves visual perception. *Neuroimage* **150**, 318-328 (2017).
125. Friese, U., Daume, J., Göschl, F., König, P., Wang, P. & Engel, A. K. Oscillatory brain activity during multisensory attention reflects activation, disinhibition, and cognitive control. *Sci. Rep.* **6**, 32775 (2016).
126. Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H. & Foxe, J. J. Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *J. Neurosci.* **31**, 18556-18567 (2011).
Human intracranial recording study which demonstrates a multi-timescale phase-amplitude coupling mechanism that supports multisensory attention.
127. Keil, J., Pomper, U. & Senkowski, D. Distinct patterns of local oscillatory activity and functional connectivity underlie intersensory attention and temporal prediction. *Cortex* **74**, 277-288 (2016).
128. Zuo, Y., Huang, Y., Wu, D., Wang, Q. & Wang, Z. Spike phase shift relative to beta oscillations mediates modality selection. *Cereb. Cortex* **30**, 5431-5448 (2020).
129. Panichello, M. F. & Buschman, T. J. Shared mechanisms underlie the control of working memory and attention. *Nature* **592**, 601-605 (2021).
130. Quak, M., London, R. E. & Talsma, D. A multisensory perspective of working memory. *Front. Hum. Neurosci.* **9**, 197 (2015).
131. Fiebelkorn, I. C. & Kastner, S. A rhythmic theory of attention. *Trends Cogn. Sci.* **23**, 87-101 (2019).
132. Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E. & Fell, J. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc. Natl. Acad. Sci. USA* **107**, 3228-3233 (2010).
133. Xie, Y., Li, Y., Duan, H., Xu, X., Zhang, W. & Fang, P. Theta oscillations and source connectivity during complex audiovisual object encoding in working memory. *Front. Hum. Neurosci.* **15**, 614950 (2021).
134. Michail, G., Senkowski, D., Niedeggen, M. & Keil, J. Memory load alters perception-related neural oscillations during multisensory integration. *J. Neurosci.* **41**, 1505-1515 (2021).
This EEG study shows that interactions between working memory and multisensory perception are expressed in multi-timescale interactions of neural oscillations.
135. Spitzer, B. & Blankenburg, F. Supramodal parametric working memory processing in humans. *J. Neurosci.* **32**, 3287-3295 (2012).
136. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127-138 (2010).
137. Arnal, L. H. & Giraud, A.-L. Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* **16**, 390-398 (2012).
138. Bastos, A. M., Lundqvist, M., Waite, A. S., Kopell, N. & Miller, E. K. Layer and rhythm specificity for predictive routing. *Proc. Natl. Acad. Sci. USA* **117**, 31459-31469 (2020).
139. Fernandez-Ruiz, A., Sirota, A., Lopes-dos-Santos, V. & Dupret, D. Over and above frequency: gamma oscillations as units of neural circuit operations. *Neuron* **111**, 936-953 (2023).

140. Arnal, L. H., Doelling, K. B. & Poeppel, D. Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex* **25**, 3077-3085 (2015).
141. Luo, H., Liu, Z. & Poeppel, D. Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biol.* **8**, e1000445 (2010).
142. Schroeder, C. E. & Lakatos, P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* **32**, 9-18 (2009).
143. van Atteveldt, N., Murray, M. M., Thut, G. & Schroeder, C. E. Multisensory integration: flexible use of general operations. *Neuron* **81**, 1240-1253 (2014).
144. Roopun, A. K. et al. Temporal interactions between cortical rhythms. *Front. Neurosci.* **2**, 145-154 (2008).
145. Chandrasekaran, C. Computational principles and models of multisensory integration. *Curr. Opin. Neurobiol.* **43**, 25-34 (2017).
146. Shams, L. & Beierholm, U. Bayesian causal inference: a unifying neuroscience theory. *Neurosci. Biobehav. Rev.* **137**, 104619 (2022).
147. Parise, C. V. & Ernst, M. O. Correlation detection as a general mechanism for multisensory integration. *Nat. Commun.* **7**, 11543 (2016).
148. Pesnot Lerousseau, J., Parise, C. V., Ernst, M. O. & van Wassenhove, V. Multisensory correlation computations in the human brain identified by a time-resolved encoding model. *Nat. Commun.* **13**, 2489 (2022).
149. Kramer, M. A., Roopun, A. K., Carracedo, L. M., Traub, R. D., Whittington, M. A. & Kopell, N. J. Rhythm generation through period concatenation in rat somatosensory cortex. *PLoS Comput. Biol.* **4**, e1000169 (2008).
150. Gelastopoulos, A., Whittington, M. A. & Kopell, N. J. Parietal low beta rhythm provides a dynamical substrate for a working memory buffer. *Proc. Natl. Acad. Sci. USA* **116**, 16613-16620 (2019).
151. Schirner, M., McIntosh, A. R., Jirsa, V., Deco, G. & Ritter, P. Inferring multi-scale neural mechanisms with brain network modelling. *Elife* **7**, (2018).
This article presents a computational modeling approach for the analysis of neural activity from humans and animals on multiple spatial and temporal scales.
152. D'Angelo, E. & Jirsa, V. The quest for multiscale brain modeling. *Trends Neurosci.* **45**, 777-790 (2022).
153. Einevoll, G. T., Kayser, C., Logothetis, N. K. & Panzeri, S. Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat. Rev. Neurosci.* **14**, 770-785 (2013).
154. Nasr, K., Haslacher, D., Dayan, E., Censor, N., Cohen, L. G. & Soekadar, S. R. Breaking the boundaries of interacting with the human brain using adaptive closed-loop stimulation. *Progr. Neurobiol.* **216**, 102311 (2022).
155. Haslacher, D. et al. In vivo phase-dependent enhancement and suppression of human brain oscillations by transcranial alternating current stimulation (tACS). *Neuroimage* **275**, 120187 (2023).
156. Fenno, L., Yizhar, O. & Deisseroth, K. The development and application of optogenetics. *Annu. Rev. Neurosci.* **34**, 389-412 (2011).
157. Grosenick, L., Marshel, J. H. & Deisseroth, K. Closed-loop and activity-guided optogenetic control. *Neuron* **86**, 106-139 (2015).
158. Zhou, H.-Y., Cai, X.-l., Weigl, M., Bang, P., Cheung, E. F. C. & Chan, R. C. K. Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **86**, 66-76 (2018).
159. Gröhn, C., Norgren, E. & Eriksson, L. A systematic review of the neural correlates of multisensory integration in schizophrenia. *Schizophr. Res. Cogn.* **27**, 100219 (2022).
160. Feldman, J. I., Dunham, K., Cassidy, M., Wallace, M. T., Liu, Y. & Woynaroski, T. G. Audiovisual multisensory integration in individuals with autism spectrum disorder: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **95**, 220-234 (2018).

161. Wallace, M. T., Woynaroski, T. G. & Stevenson, R. A. Multisensory integration as a window into orderly and disrupted cognition and communication. *Annu. Rev. Psychol.* **71**, 193-219 (2020).
162. van Laarhoven, T., Stekelenburg, J. J., Eussen, M. L. & Vroomen, J. Atypical visual-auditory predictive coding in autism spectrum disorder: electrophysiological evidence from stimulus omissions. *Autism* **24**, 1849-1859 (2020).
163. Hirano, S. et al. Phase-amplitude coupling of the electroencephalogram in the auditory cortex in schizophrenia. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **3**, 69-76 (2018).
164. Moran, J. K., Michail, G., Heinz, A., Keil, J. & Senkowski, D. Long-range temporal correlations in resting state beta oscillations are reduced in schizophrenia. *Front Psychiatry* **10**, 517 (2019).
165. Cruz, G., Grent-'t-Jong, T., Krishnadas, R., Palva, J. M., Palva, S. & Uhlhaas, P. J. Long range temporal correlations (LRTCs) in MEG-data during emerging psychosis: relationship to symptoms, medication-status and clinical trajectory. *Neuroimage Clin.* **31**, 102722 (2021).
166. Rognini, G. et al. Multisensory bionic limb to achieve prosthesis embodiment and reduce distorted phantom limb perceptions. *J. Neurol. Neurosurg. Psychiatry* **90**, 833-836 (2019).
167. Aurucci, G. V., Preatoni, G., Damiani, A. & Raspopovic, S. Brain-computer interface to deliver individualized multisensory intervention for neuropathic pain. *Neurotherapeutics* **20**, 1316-1329 (2023).
168. Lloyd-Esenkaya, T., Lloyd-Esenkaya, V., O'Neill, E. & Proulx, M. J. Multisensory inclusive design with sensory substitution. *Cogn. Res. Princ. Implic.* **5**, 37 (2020).
169. Mackin, A. & Bull, D. Characterizing the spatiotemporal envelope of the human visual system through the visibility of temporal aliasing artifacts. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **37**, 1116-1127 (2020).
170. Mori, S., Kikuchi, Y., Hirose, N., Lepage, H. & Wong, W. Auditory gap detection: psychometric functions and insights into the underlying neural activity. *Biol. Cybern.* **112**, 575-584 (2018).
171. Victor, J. D. How the brain uses time to represent and process visual information. *Brain Res* **886**, 33-46 (2000).
172. Uchida, N., Poo, C. & Haddad, R. Coding and transformations in the olfactory system. *Annu. Rev. Neurosci.* **37**, 363-385 (2014).
173. Kayser, C., Montemurro, M. A., Logothetis, N. K. & Panzeri, S. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* **61**, 597-608 (2009).
174. Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H. & Wang, X.-J. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* **88**, 419-431 (2015).
175. Vroomen, J. & Keetels, M. Perception of intersensory synchrony: a tutorial review. *Atten. Percept. Psychophys.* **72**, 871-884 (2010).
176. Manger, P. R., Engler, G., Moll, C. K. E. & Engel, A. K. The anterior ectosylvian visual area of the ferret: a homologue for an enigmatic visual cortical area of the cat? *Eur. J. Neurosci.* **22**, 706-714 (2005).
177. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **12**, 97-136 (1980).
178. Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* **94**, 115-147 (1987).
179. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193-222 (1995).
180. Johansen-Berg, H. & Lloyd, D. M. The physiology and psychology of selective attention to touch. *Front. Biosci. Landmark* **5**, 894-904 (2000).
181. Fritz, J. B., Elhilali, M., David, S. V. & Shamma, S. A. Auditory attention – focusing the searchlight on sound. *Curr. Opin. Neurobiol.* **17**, 437-455 (2007).

182. Freeman, W. J. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biol. Cybern.* **56**, 139-150 (1987).
183. Gray, C. M. & Singer, W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* **86**, 1698-1702 (1989).
184. Gray, C. M., König, P., Engel, A. K. & Singer, W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* **338**, 334-337 (1989).
185. Fries, P., Reynolds, J. H., Rorie, A. E. & Desimone, R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* **291**, 1560-1563 (2001).
186. Ghazanfar, A. A. & Schroeder, C. E. Is neocortex essentially multisensory? *Trends Cogn. Sci.* **10**, 278-285 (2006).
187. Voloh, B. & Womelsdorf, T. A role of phase-resetting in coordinating large scale neural networks during attention and goal-directed behavior. *Front. Syst. Neurosci.* **10**, 18 (2016).
188. Gingras, G., Rowland, B. A. & Stein, B. E. The differing impact of multisensory and unisensory integration on behavior. *J. Neurosci.* **29**, 4897-4902 (2009).

Glossary

Oscillation
Rhythmic pattern of electrical activity generated by a cooperating group of neurons.

Time windows
Temporal epochs in which significant information about sensory stimuli can be encoded or processed; such windows can vary in duration, giving rise to different processing timescales.

Autocorrelation
Correlation between values of a signal at different time points, as a function of the lag between these time points.

Intrinsic neural timescales
Processing time windows inferred from the autocorrelation of neural signals in a given brain region.

Intrinsic coupling modes
Neural coupling patterns that are not imposed by external factors but are generated in the brain.

Cross-frequency coupling
Coupling of the power and/or phase of neural oscillations across different frequency bands.

Amplitude modulation
Change of the amplitude of a neural oscillation in response to an external input.

Phase resetting
Shift of the phase of a neural oscillation in response to external input.

Entrainment
Synchronization of neural activity to external rhythmic input.

Phase coupling
Dynamic coupling based on the correlation of the phase of neural signals.

Envelope coupling
Dynamic coupling based on the correlation of the amplitude envelope of neural signals.

Phase-amplitude coupling
Modulation of the amplitude of an oscillatory signal by the phase of another oscillatory signal.

Power
Square of the amplitude of a sinusoidal wave.

Crossmodal
An interaction between neural systems in which input from one sensory modality influences the processing in another sensory modality.

Functional connectivity
Statistical relationship between signals recorded from different neurons or brain regions.

Double flash illusion
Illusory perception of two flashes when a single flash is presented together with two auditory clicks or two tactile stimuli.

McGurk effect
Fusion of an auditory syllable or word paired with a conflicting visual syllable or word into a combined multisensory percept.

Causal prior
An person’s belief about the causal structure of a multisensory stimulus.

Scale-free dynamics
Lack of a characteristic timescale in the dynamics of a neural process.

Long-range temporal correlations
Persistence of correlations in time series data across multiple timescales.

Power-law scaling
Relation between two variables in which one variable varies as a power of the other; this yields a linear relation if both variables are displayed in a logarithmic manner.

Criticality
State of a system close to a phase transition, in which spatiotemporal correlations are highly susceptible to perturbations.

Predictive processing
Brain processes involved in generating and updating predictions about sensory inputs or events.

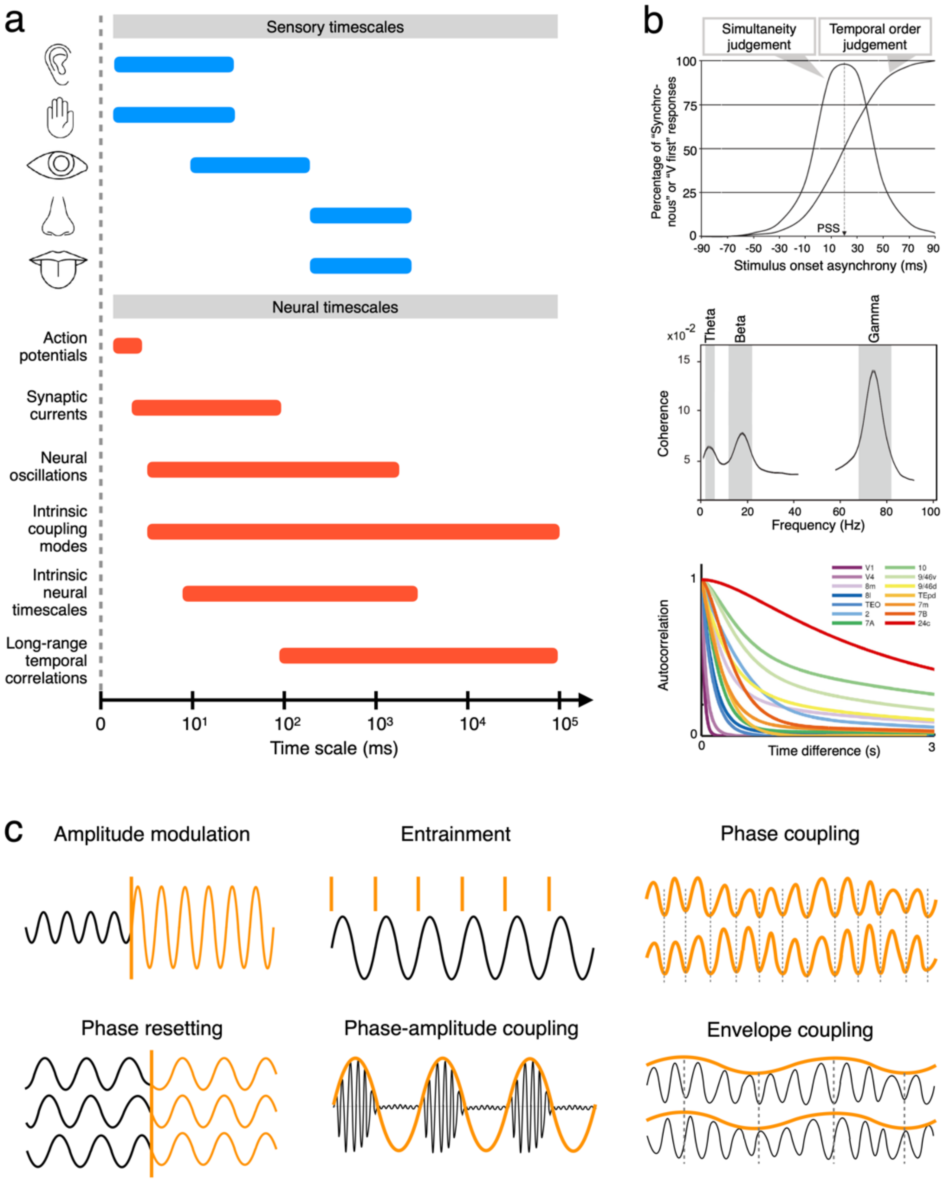

Fig. 1 | Multiple timescales and neural mechanisms relevant for multisensory processing. a | Estimates for sensory and neural timescales. Note that timescales differ considerably between the sensory systems. Ranges indicated for sensory timescales are based on behavioral evidence[169,170], neurophysiological data[171,172], and data on information encoding in sensory signals[7,173]. Neural timescales result from the duration of basic events in nerve cells, such as action potentials and synaptic currents, as well as from network dynamics. Relevant aspects of network dynamics include autocorrelation-based estimates of intrinsic neural timescales[23,174], long-range temporal correlations[111,113], neural oscillations[8,9], and intrinsic coupling modes[11,17,26]. Note that network-based timescale measures can vary over several orders of magnitudes. **b** | Evidence for perceptual and neural timescales. (Top) Behavioral quantification of sensory timescales. The example illustrates data for judgement on simultaneity or temporal order of audiovisual stimuli. (Middle) Quantification of timescales by analysis of spectral properties of neural signals. Data show distinct frequency bands in which phase coupling occurs in monkey visual cortex. (Bottom) Neural timescales can also be captured by autocorrelation analysis. The panel shows results on modeling of autocorrelation functions for different areas in monkey cortex. **c** | Features of neural oscillations relevant for multisensory integration. Modulation of both phase and amplitude of oscillations can be involved in MI. Relevant mechanisms include: (left, top panel) amplitude modulation of a neural oscillation after the onset of a sensory stimulus; (left, bottom panel) resetting of the phase of ongoing oscillations by a sensory stimulus transient; (middle, top panel) entrainment of neural oscillations by rhythmic external events; (middle, bottom panel) coupling between the amplitude of a high-frequency oscillation and the phase of an oscillation in a lower frequency band; (right, top panel) coupling, or coherence, between the phase of oscillations of two neuronal populations; (right, bottom panel) correlated amplitude envelope fluctuations between different neuronal populations. Part **b** top panel adapted from[175], middle panel adapted from[28] and bottom panel adapted from[174].

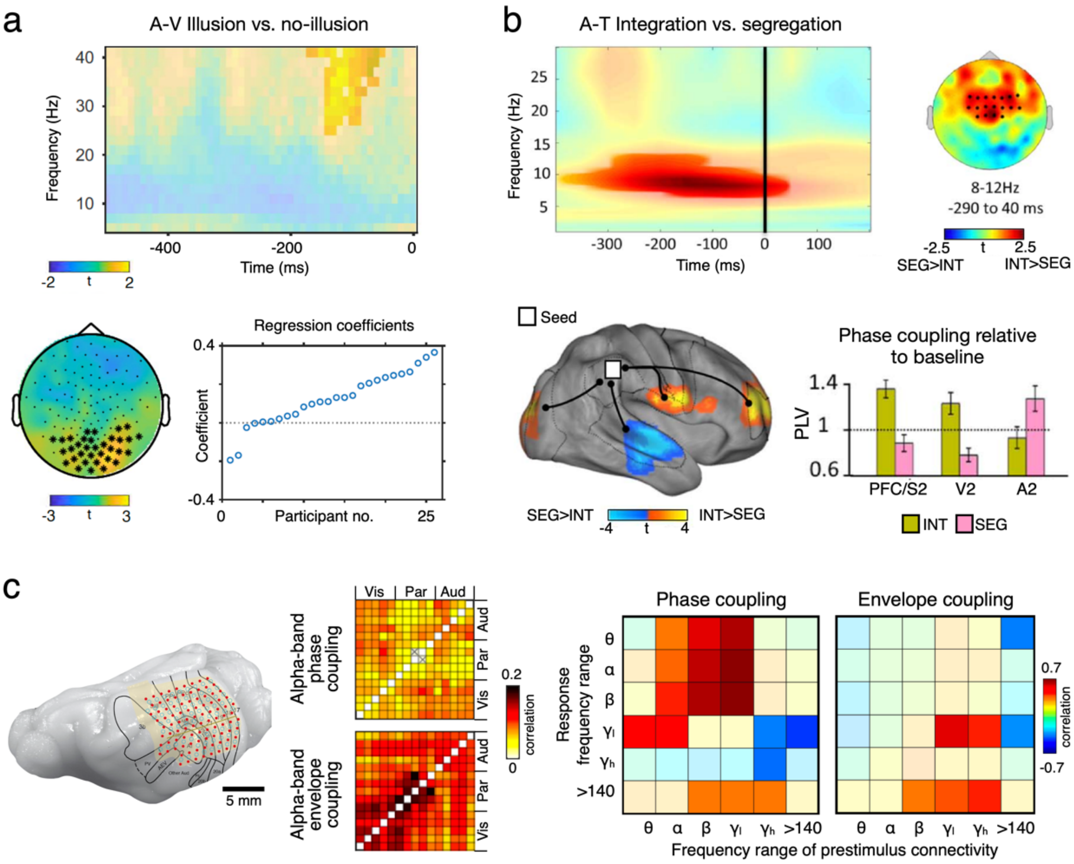

Fig. 2 | Multisensory processing and multiscale dynamics in ongoing activity. Ongoing cortical dynamics, as reflected in phase and envelope coupling as well as power and phase of prestimulus oscillations at different frequencies, predicts upcoming multisensory integration. **a** | Prestimulus beta band power predicts perception in the sound-induced flash illusion. Power in the high beta and low gamma range was higher before stimulus onset (at 0 ms) in trials where the auditory-visual (A-V) illusion was perceived (top panel). This effect was significant over occipital and parietal channels (asterisks in bottom panel, left). The strength of the effect varied across participants, as indicated by the distribution of single-trial regression coefficients (bottom panel, right). **b** | Prestimulus alpha band power and phase coupling predict integration (INT) or segregation (SEG) of auditory-tactile (A-T) stimuli. Alpha power was higher in trials where participants perceived an auditory and a tactile stimulus as belonging to the same event (top panel). Multisensory integration was also predicted by prestimulus coupling between intraparietal cortex (seed), somatosensory (S2), prefrontal (PFC), visual (V2) and auditory (A2) cortex. **c** | Prestimulus network coupling predicts power changes during processing of multisensory stimuli in ferret cortex. (Left) Data were recorded using micro-electrocorticographic arrays from visual (Vis), auditory (Aud) and parietal (Par) areas in anesthetized ferrets. Red dots indicate electrode contacts. (Middle) Patterns of phase (top) and envelope (bottom) coupling differ in their topographies. Example shown here depicts alpha band coupling. (Right) Coupling in pre-stimulus epochs predicts multisensory processing. The matrices show the correlation between pre-stimulus phase and envelope coupling and multisensory responses. Multisensory responses were quantified as stimulus-induced power changes in different frequency bands. Note that the two types of coupling differ in their relation to response processing. Part **a** adapted from[89], part **b** adapted from[92], and part **c** adapted from[176] (left) and[42] (middle, right), respectively.

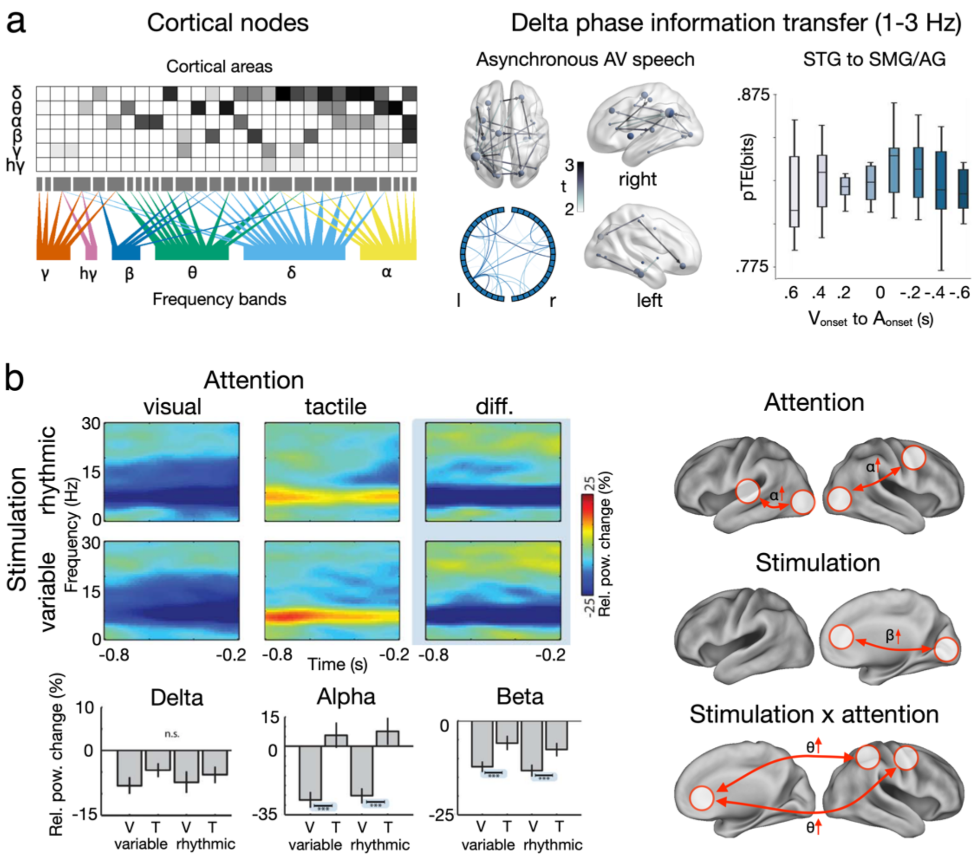

Fig. 3 | Multiscale dynamics during cognitive processing. Multisensory processing of speech and intersensory attention are reflected in nested hierarchical networks, power modulation and functional coupling of neural oscillations at different frequencies. **a** | (Left) Hierarchically nested functional networks with different time scales during audiovisual (AV) speech processing. (Top panel) Darker squares indicate a higher strength of the respective node in frequency-specific networks. (Bottom panel) High-frequency networks are spatially nested into low-frequency networks during AV speech processing. (Middle) Illustration of delta phase information transfer in a left hemisphere dominated cortical network, in which the left supramarginal gyrus/angular gyrus (SMG/AG) formed a central hub. The size of the node symbols reflects their strength in the network. (Right) Information flow from the superior temporal gyrus (STG) to SMG/AG rapidly adapts to temporal asynchronies between visual and auditory speech onsets. **b** | Influence of attention and stimulus rhythmicity on cortical network dynamics. (Left) Power of ongoing oscillations in the visual cortex during an intersensory visuo-tactile attentional cueing task. Attention to upcoming visual (V) compared to tactile (T) stimuli is associated with greater alpha and beta band desynchronization, regardless of whether the stimuli are presented in a rhythmic or variable temporal fashion. (Right) These effects are paralleled by enhanced functional coupling in the alpha band (top panel) between visual cortex and temporal and somatosensory cortices when tactile compared to visual stimuli are attended. In addition, variable stimulation is reflected in a stronger beta band functional coupling (middle panel) between visual and frontal cortex. Finally, interactions between stimulus rhythmicity and intersensory attention are reflected in theta band modulations of functional connectivity (bottom panel) in a cortical network involving frontal, posterior parietal and somatosensory regions. Part **a** adapted from[122], part **b** adapted from[37] (left) and[127] (right).

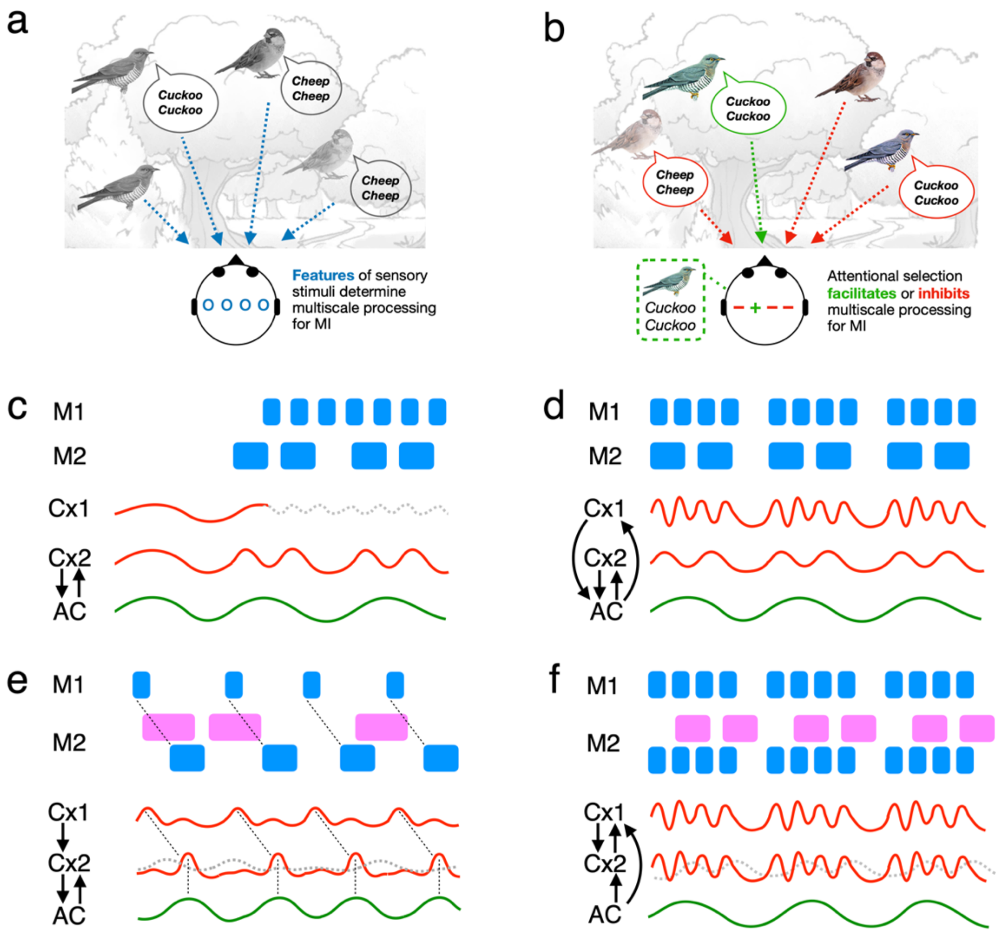

Fig. 4 | Scenarios for multi-timescale MI. a | Stimulus scenario with entrainment of multisensory processing driven bottom-up by temporal structures in the stimuli. **b** | Stimulus scenario with predominance of internally generated temporal patterns, reflecting attentional mechanisms that lead to selection or suppression of multisensory stimuli. **c** | Multi-timescale dynamics can support multisensory segregation. In this example, top-down influences mediated through cross-frequency coupling between higher-order cortex and Cx2 lead to selection of input stream M2 and its segregation from input M1. The scenario also illustrates that priors contained in slow oscillations in ongoing activity could facilitate the selection of an input stream (M2) that matches, through cross-frequency coupling, with the phase of the ongoing oscillation. **d** | Multi-timescale processing would allow to couple sensory inputs that differ in fast temporal components by phase-amplitude coupling or envelope coupling to slower rhythms in higher-order association cortex. Here, cross-frequency coupling facilitates the information flow between early and higher order cortex. **e** | Multisensory integration through crossmodal phase resetting. Here, the M1 stimulus predicts only one of the two input patterns in modality M2. The crossmodal phase resetting in Cx2 leads to a temporal pattern (red curve) that is also aligned with a slow oscillation in higher-order cortex; this does not hold for the temporal pattern induced by the competing stimulus in Cx2 (dashed grey curve). **f** | Selection of two competing stimuli (in Cx2) can occur by coupling to temporal patterns in another modality (Cx1), and by top-down influences from temporal patterns in higher-order cortex. The oscillatory pattern induced by the second stimulus in M2 (dashed grey curve in Cx2) if effectively suppressed due to the misalignment with the dominant multi-timescale pattern. Conventions for panels c-f: M1, M2 – input streams in different sensory modalities; Cx1, Cx2 – cortical areas involved in processing of the unisensory inputs; AC – supramodal association cortex; arrows indicate dominant directions of information transfer for the respective scenario.

Box 1 – Multisensory versus unisensory processing

Integration, segregation, hierarchical structuring and selection of information are not only relevant in multisensory processing but are also critical for unisensory processing. These aspects have probably been best studied for the visual system, particularly in the context of visual feature binding[177,178] and visual attention[179]. However, they are also well investigated for other sensory modalities[180,181]. The fundamental role of fast neuronal oscillations for information processing was first discovered in the olfactory system[182] and in the visual system[183], showing specific relations of the power of neural oscillations as well as phase coupling of oscillatory signals across neural populations to sensory processing and bottom-up integration of information[184]. Further studies revealed a relevance of power and phase modulations in top-down control of unisensory processing, e.g., in attentional selection[185] and prediction of stimuli[11]. While the basic requirements for integration, segregation, structuring and selection may be comparable for uni- and multisensory processing, and the same basic oscillatory mechanisms seem to be in place, several differences are notable. First, the brain networks and the neural populations involved in multisensory processing differ from those for unisensory processing. The former involves neurons with multisensory response profiles, e.g., in the superior colliculus, that show enhanced (super-additive) firing rates in response to crossmodal stimuli compared to the linear combination of the corresponding unisensory stimuli[3]. Similarly, neural populations in superior temporal, intraparietal and prefrontal association areas show specific multisensory response profiles[143,186]. Second, in case of multisensory processing the oscillatory multi-timescale mechanisms have to operate across different subsystems of the brain in order to serve for coupling, entrainment or phase-resetting of signals between different modalities. For instance, the role of crossmodal phase resetting (FIG. 4e) in the alignment and augmentation of responses between different modalities very likely differs from the function of phase resetting in the unimodal case[30,69,187]. Third, the computations performed by the networks involved seem to differ between the multisensory and the unisensory case, e.g., with respect to stimulus prediction and error correction. One important aspect is that cross-modal processing provides independent samples of features of the environment through different sensory channels, whereas signals within the same modality can exhibit substantial covariance; therefore, multisensory integration can yield better error reduction than unisensory integration[123,188].