Review

Rough is salient: a conserved vocal niche to hijack the brain's salience system

Luc H. Arnal and Noémi Gonçalves

Université Paris Cité, Institut Pasteur, AP-HP, INSERM, CNRS, Fondation Pour l'Audition, Institut de l'Audition, IHU reConnect, Paris 75012, France

LHA, 0000-0002-2226-6497

The propensity to communicate extreme emotional states and arousal through salient, non-referential vocalizations is ubiquitous among mammals and beyond. Screams, whether intended to warn conspecifics or deter aggressors, require a rapid increase of air influx through vocal folds to induce nonlinear distortions of the signal. These distortions contain salient, temporally patterned acoustic features in a restricted range of the audible spectrum. These features may have a biological significance, triggering fast behavioural responses in the receivers. We present converging neurophysiological and behavioural evidence from humans and animals supporting that the properties emerging from nonlinear vocal phenomena are ideally adapted to induce efficient sensory, emotional and behavioural responses. We argue that these fast temporal —rough—modulations are unlikely to be an epiphenomenon of vocal production but rather the result of selective evolutionary pressure on vocal warning signals to promote efficient communication. In this view, rough features may have been selected and conserved as an acoustic trait to recruit ancestral sensory salience pathways and elicit optimal reactions in the receiver. By exploring the impact of rough vocalizations at the receiver's end, we review the perceptual, behavioural and neural factors that may have shaped these signals to evolve as powerful communication tools.

This article is part of the theme issue 'Nonlinear phenomena in vertebrate vocalizations: mechanisms and communicative functions'.

1. An essential and universal vocalization

Vocal signals play an essential role in alerting conspecifics and promoting survival [1,2] by effectively conveying warnings, especially in environments where visibility is limited. One common strategy to intensify exogenous sensory perception and ensure powerful reactions in the listener is screaming. This loud vocal response to fear or distress stands out as one of the most potent vocalizations that one can produce—reflexively or voluntarily [3]—to trigger a maximal response in the receivers' brain, as well as fast and efficient reactions [4,5]. While speaking and singing require input and training, screaming holds an immediate and innate position in our communications toolkit [6]: one shared by numerous species exhibiting comparable vocal properties and behavioural consequences [7–9].

The human newborn's immediate and instinctual emission of a cry upon birth, devoid of prior rehearsal, not only clears airways but also serves as a fundamental communication signal [10]. The acoustic features of a baby's cry can provide valuable information about their health state, making the newborn's cry not just a sign of life but also an interesting diagnostic tool [11,12]. Being born at an immature stage, human neonates are totally dependent on the care of their parents and rely on their capacity to capture their attention using loud utterances such as screams [13]. Screams possess an inherent ability to grab attention and are very difficult to ignore. They carry

ecologically relevant attributes and are notably salient, thereby standing to ensue rapid reaction, requisitioning attentional resources with priority over other inputs [4,5].

In humans, screams occupy a distinctive position within the vocal repertoire, rarely used in routine communication [4] and inappropriate for most adult social situations in Western cultures. Screams extract a high cost both from the speaker and the listener. In the speaker, they require high energy outputs, causing strain in the vocal folds as well as throat pain [14]. In the listener, they can evoke high arousal [15–17], emotional distress and anxiety [18]. Consequently, speakers typically scream only in moments of severe distress, using this vocalization sparingly to justify the associated costs to themselves and others in their social group. These physical and societal pressures arguably encourage the parsimonious use of such high-emergency signals.

Alarm vocalizations cover a range of vocalizations beyond screams, including various acoustic features (e.g. pitch and timbre) to convey affective [19] or semantic information [20]. Here, we mainly focus on primordial, non-referential vocalizations that are intended to capture attention and warn or deter aggressors [4]. In humans, we specifically define 'screams' as loud, salient and involuntary vocalizations that are typically characterized by vocal harshness. While the term 'scream' is scarcely used in other species, this definition intuitively applies to animal vocalizations emitted in response to highly urgent situations. This distinction separates screams from other vocalizations, such as cries, whines or roars. While cries, for instance, often signal distress or need, particularly in infants, alarm vocal signals like screams stand out for their intense urgency, aimed at alerting conspecifics to immediate danger rather than communicating a specific emotional state. Focusing on alarm-like screams sets them apart from other rough vocalizations—such as joyful or excited screams—that, while sharing acoustic roughness, serve different communicative purposes and often express positive arousal rather than distress [21]. Accordingly, this review centres exclusively on alarm and distress calls, excluding other emotionally expressive, non-alarm signals that do not function as urgent warnings. Given the instinctual and primordial nature of alarm vocal signals, we propose that these vocalizations could be regarded as a shared precursor of vocal signalling, suggesting that identifying inherently salient or arousing aspects of such inputs may enhance our comprehension of the evolution of perception and communication systems. We review the perceptual, behavioural and neural constraints that may have shaped these vocalizations to evolve into powerful tools for effective alarm communication. We argue that a deeper understanding of the perceptual, neural and behavioural mechanisms that underlie salience processing in humans and animals may illuminate the processes that have contributed to the preservation of these acoustic features in vocal alarm communication.

2. Vocal production

In mammals, vocalizations primarily result from the vibration of the vocal folds in the larynx [22], which are then filtered through the vocal tract [23]. While more typical vocalizations are driven by smooth harmonic changes in the larynx, more abrupt changes and tightening can cause chaotic dynamics leading to alert and distress calls, also referred to as nonlinear phenomena (NLP). The resulting sound possesses salient and arousing acoustic features efficiently acting upon sensory, emotional, cognitive and behavioural levels [24–26]. In mammals, NLP are more present in alarm and distress calls [27] and screams [4] than other vocalizations, serving as an effective tool to grab attention, to seek help from conspecifics [28] or to discourage aggressors [29] while preventing habituation [30]. The presence of NLP extends beyond adults; juveniles also produce similar vocalizations (non-humans [31–33], see also [34]; humans [35–37]), supporting the innate nature of such vocal behaviour in eliciting analogous responses [38]. While anatomical structures for vocal production can vary among animals, the use of nonlinear features for communicative purposes is widespread across the animal kingdom, extending beyond mammals (see, e.g. this special issue and following references on mammals: humans [4,39], marmots [40], also see [41], elephants [42], dogs [43,44], meerkats [45], manatees [46], chimpanzees [47], monkeys [48], koalas [49], giant pandas [50], altai pikas [51], sambars and muntjacs [52], wapitis [53], cows [54], red wolf [55], marine mammals: whales [56–58], dolphins [59], as well as birds such as sparrows [60], see also [61] and amphibians such as frogs [62]), suggesting an evolutionary continuum of warning communication across species (see examples in figure 1). Interestingly, cross-species studies have shown that mammalian species are capable of recognizing urgency in calls from other species, underlining the conservation of arousal cues in vocal communication. Humans have the capacity to accurately decode emotional perception of primate vocalizations [63–65], a capacity that even extends to non-mammalian vertebrates [66]. Even some reptiles, such as crocodiles, can react to vocal emotions conveyed by bonobo, chimpanzee and human alarm calls and are attracted to infant cries, although it is not clear whether this effect pertains to the detection of nonlinear cues [67]. Altogether, these observations suggest that despite diversity in vocal production mechanisms, many animals tend to utilize and produce similar acoustic features to be salient (see examples in figure 1B). The selection of nonlinear acoustic features for communication may be constrained by the receiving rather than the producing end, i.e. it may be optimized to align with the sensory processing constraints of the receivers.

3. Acoustic characterization

Vocal alarm signals have arguably emerged to maximize salience to ensure unconditional detection by the receiver. Therefore, analysing the acoustic structure of such sounds can reveal how natural communication has shaped vocal signals to maximize exogenous attention within the receivers' sensory system. Among the numerous analytical methods used to measure NLP, the modulation power spectrum (MPS, see figure 1A, right panel) is particularly convenient in the current context—and more generally in auditory neuroscience—as it provides a neurally and ecologically relevant parametrization of sounds [68–71]. The MPS quantifies the power in temporal and spectral modulations by calculating the two-dimensional Fourier transform
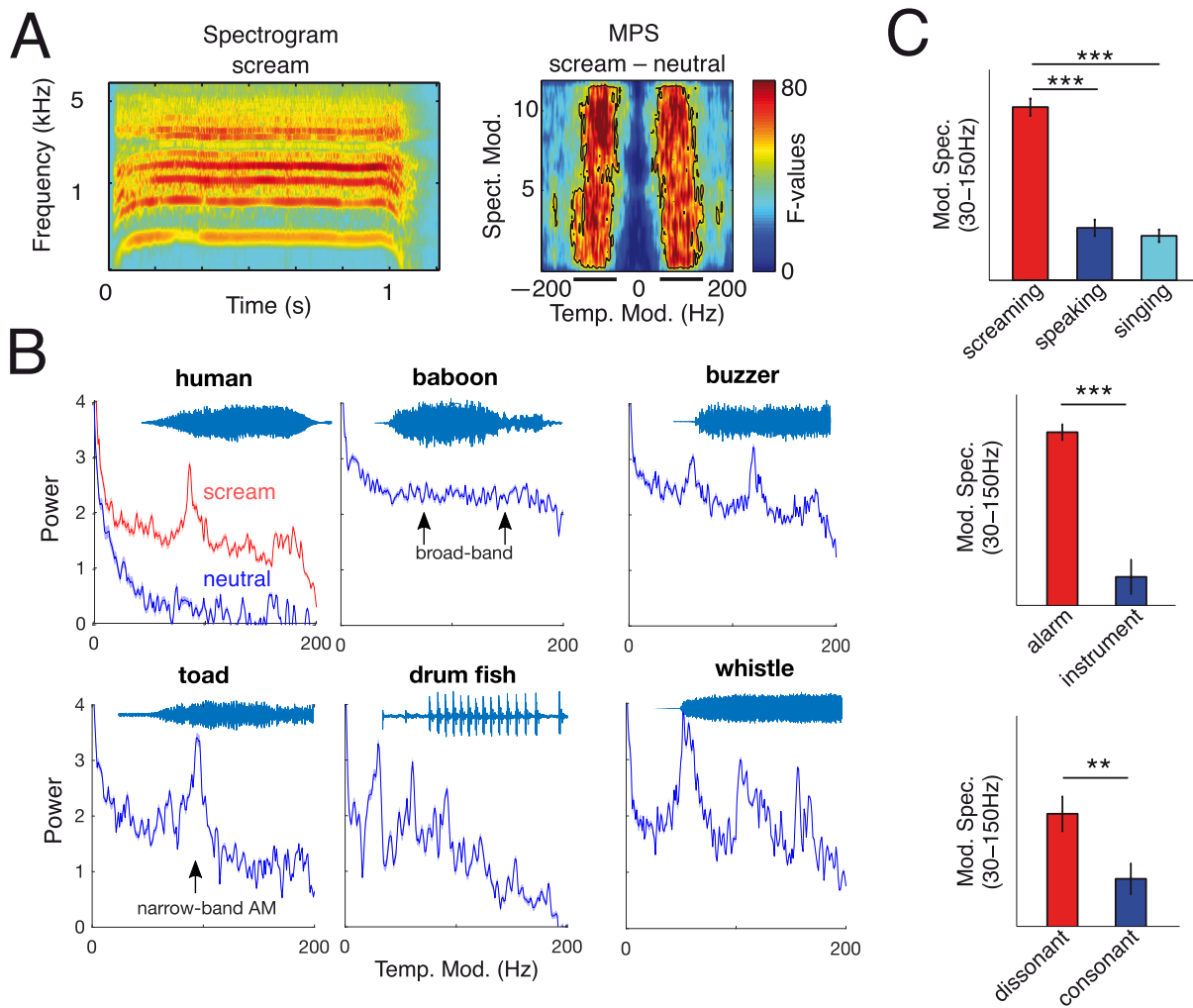
Figure 1. Acoustic properties of salient vocal signals across species. (A) Screamed vowel [a] spectrogram from one participant (top left); human scream vocalizations exploit a specific and isolated acoustic niche in the modulation power spectrum (MPS) (right) [4]. (B) Examples of temporal modulations in mammal (humans and baboons) screams, amphibian (toad) and fish (drum fish) vocalizations and artificial sounds (buzzer and whistle). L.H. Arnal, 2019, unpublished data. (C) MPS roughness. Top: screams (red), natural speech (dark blue) and musical vocalizations (a cappella) Centre: artificial alarm sounds versus musical instruments. Bottom: dissonant versus consonant sounds. Screams, compared with speaking and singing vocalizations, exhibit higher MPS values (e.g. average magnitude of temporal modulations in the roughness range (30–150 Hz). MPS values are also stronger in alarm and dissonant sounds compared with musical and consonant sounds, respectively. Roughness is also exploited in artificial alarm signals. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Error bars indicate the s.e.m. Spect. Mod.: spectral modulations, Temp. Mod.: temporal modulations. Adapted from Arnal *et al.* [4].

of a spectrogram (see figure 1A). NLP generally create distortions in the acoustic signal (e.g. chaotic modulations) that are visible on a spectrogram (subharmonics, chaos, frequency jumps) and can be quantified using the MPS. Focusing on human screams, we previously demonstrated that screams—compared to neutral vocalizations—enhance modulations in a specific acoustic/perceptual regime. This roughness range (see figure 1A, right panel) [4] reflects a restricted amplitude modulation (AM) frequency range [30–150 Hz] of the audible spectrum, perceived as salient and aversive, that has been preserved to communicate distress or danger. Interestingly, this special acoustic regime, 'roughness' appears to be almost exclusively used for alarm communication (figure 1B,C) but only minimally in neutral vocalizations such as speech or singing (figure 1C). Taken together, this supports the notion that screams—but not neutral vocalizations—occupy a specific acoustic niche in the audible spectrum—the roughness range—that is preserved to communicate high arousal (e.g. owing to danger or distress), emphasizing the biological and evolutionary relevance of such modulations. Many other animals (which generally have a more limited vocal repertoire than humans) appear to adopt this approach as well, exploiting rough cues to express urgency (see examples in figure 1B and above). Overall, sounds (whether natural or artificial) that aim to be salient often exhibit broadband spectral energy in the roughness range, sometimes combined with narrower-band amplitude modulations within the same frequency range (see figure 1B).

4. Roughness in culture

Auditory roughness was initially defined by Hermann von Helmholtz as the perception when two sounds that are tonally close are heard simultaneously [72], linking this attribute to musical dissonance [73,74]. Although initial characterization of roughness perception arose in psychoacoustic studies focusing on the comprehension of musical dissonance, a wealth of

recent studies has supported the ecological and emotional relevance of this feature in our everyday lives. Roughness has been shown to be a critical aspect of music perception [75] as well as defining orchestral timbres [74]. Music composers often use non-harmonic distortions to influence our emotional responses [76,77]. Frightening scenes in films are often carried by 'scream-like' music (e.g. scratchy violins) exhibiting roughness and correlating with negative valence and higher arousal levels in spectators [17]. While it is tempting to suggest that the use of roughness is shared across cultures, its deployment may vary across cultures [78]. Existing research is largely limited to WEIRD populations, and further studies are needed to understand cross-cultural differences and universal aspects of roughness perception.

Beyond natural vocalizations, roughness is also exploited in sirens, fire alarms and medical alerts [4,79], suggesting that alarm sound designers empirically converge on choosing this attribute to enhance signals' salience and to capture attention and/or convey danger. Amplitude modulations in the 30–150 Hz frequency range generate a rough sonic texture [80], inducing unpleasant or harsh subjective experience [76] commensurate with the power of such modulations [4]. Comparing the roughness of non-vocal sounds, we have shown that artificial alarms and dissonant intervals present stronger amplitude modulations in the roughness range than pitch-matched musical instruments and consonant sounds, respectively (figure 1C, middle and bottom panels). Beyond the auditory modality, roughness appears to be a fundamental sensory property that is used across different modalities [81]: strobe lights and tactile vibrators, which arguably have warning functions comparable to auditory buzzers—namely to capture attention—appear to exploit a very similar strategy to enhance perceptual salience by introducing fast temporal modulations in the sensory input [72].

5. Behavioural and affective reactions

Across artificial and natural sounds, roughness significantly impacts behavioural and emotional reactions. At the subjective level, the amount of roughness is highly predictive of participants' aversion ratings to a sound, as shown in both lab and online experiments (see figure 2A, left and right panels [4]). Roughness also enhances behavioural efficiency (performance and speed) to spatially localize vocalizations: both natural and artificial rough screams are localized more efficiently than smooth neutral voices (figure 2B), demonstrating that roughness is a sufficient ingredient to enhance salience and facilitate behavioural reactions. In a psychoacoustic experiment [5], we tested the subjective impact of sounds (click trains of constant amplitude) on human participants across the roughness range and beyond (10–250 Hz). While one should expect to observe a linear profile of aversion commensurate with increasing sound intensity as a function of frequency (red line in figure 2C), sound aversion follows a nonlinear profile and is maximal in the roughness range, specifically at 40 Hz (figure 2C). Interestingly, the nonlinear increase in the aversion profile appears to be upper-bounded by the capacity of the brain to track fast temporal modulations and vanishes beyond the 'temporal sampling limit' (>130 Hz) at which temporal cues get fused into a continuous (pitch) percept. Beyond this limit, aversive responses linearly reflect the stimulus frequency (pitch). In addition to testing the effect of roughness on sound aversion and localization performance, we further hypothesized that such cues may enhance the detection of sounds presented at a low signal-to-noise ratio. This experiment showed that fast amplitude modulations in the roughness range (but not above) enhance the detection of tone carriers presented at low-level intensity (figure 2D, L.H. Arnal, 2021, unpublished data). Overall, the data converge to show that roughness can benefit the detection of very faint sounds, supporting that such acoustic features can extend the range of warning, even at a great distance between the emitter and the receiver.

Rough sounds not only have a considerable impact on perception but also on emotions. They are almost impossible to ignore or suppress perceptually, suggesting that their processing may be prioritized in a way that prevents habituation and competition with stimuli in other modalities. In extreme cases, overexposure to rough sounds like screams causes tremendous stress in the listener's brain [83]. Persistent infant screams are often related to parental exhaustion, depression and even extreme responses sometimes leading to dramatic consequences such as shaken baby syndrome [84,85]. This also aligns with findings in rats showing that chronic exposure to screams has a considerable impact on cognitive and physiological processes, altering memory and monoamine levels, respectively [86,87].

In animals, studies show that during fear conditioning tasks, distress calls carrying rough amplitude modulations are emitted both in rats [88] and bats [89–91], arguably reflecting the animal's state of fear and intent to deter or fly away from the experimenter. In chimpanzees' social interactions, gaze direction is enhanced during screams compared with other vocalizations, suggesting that roughness may elicit arousal-mediated social engagement [92]. In the same way, parents spontaneously run to their screaming offspring and the AM tone's roughness enhances mice's attraction to a sound source as compared with sounds outside this range [93]. In humans, auditory roughness inhibits micro-saccades, highlighting the capacity of salient sounds to interrupt active visual exploration and rapidly orient attentional resources towards exogenous events [82].

6. Physiological responses

In parallel to evidence showing that rough sounds modulate subjective emotional responses, other works have investigated their influence on physiological and homeostatic body signals [94,95]. Several studies have linked stress and anxiety with cardiac changes in humans, with similar findings observed in animals [96–98]. In bats, Hechavarria et al. [89] used heart rate as an indicator of autonomic changes, showing that rough-like vocalizations accelerate heart rate potentially by stimulating neurons in the basal and central amygdala [99]. Whether this effect occurs via a direct routing from subcortical auditory primary regions to the amygdala [100] or indirectly through the auditory cortex [101] is still unclear (see below). In humans, baby screams often elicit greater heart rate reactivity responses, with an increase of skin conductance in parents, suggesting
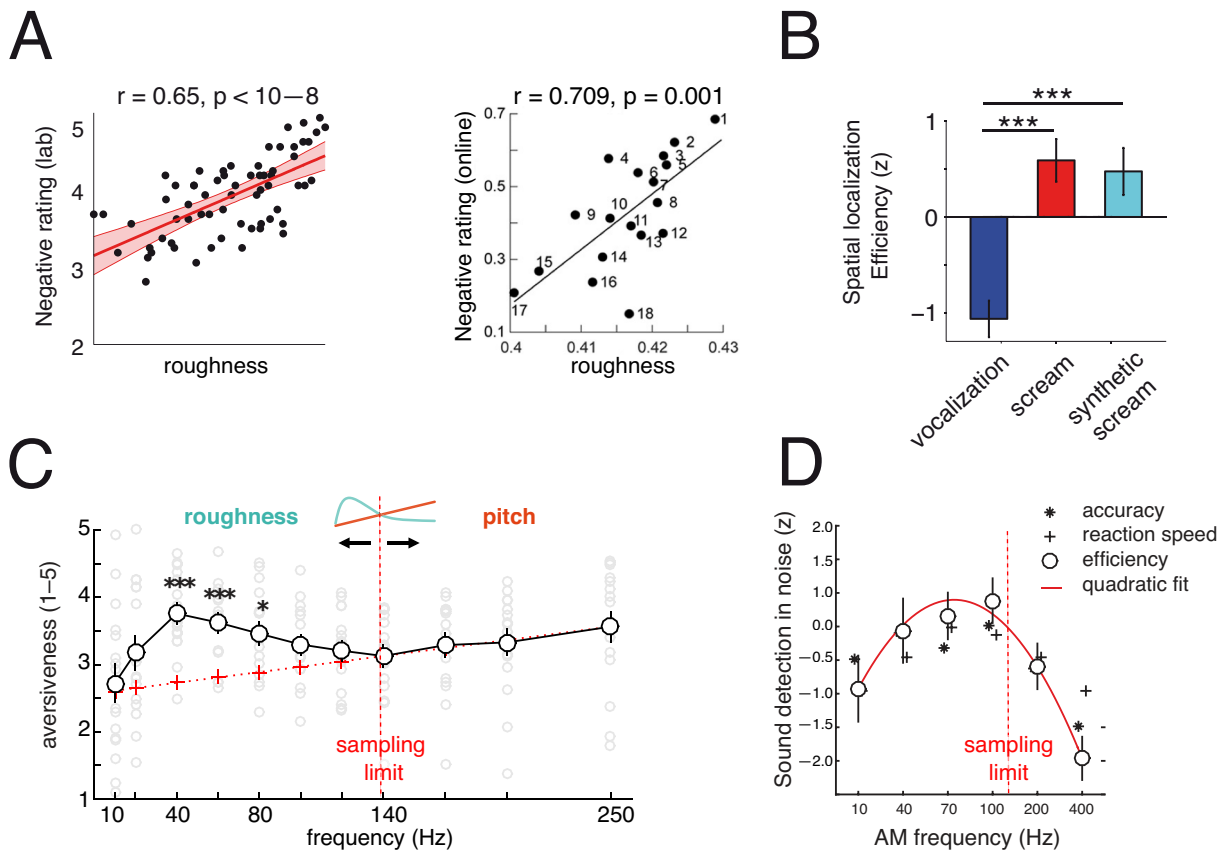
Figure 2. Affective reactions and behavioural impact of roughness. (A) Negative subjective rating increases with roughness assessed in lab conditions (right panel) and online [82]. Adapted from Zhao *et al*. [82]. (B) Roughness improves behavioural efficiency in spatially localizing rough vocalizations, whether produced naturally (screams) or artificially (synthetic screams), as compared with neutral, pitch-matched controls. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Error bars indicate the SEM. Adapted from Arnal *et al*. [4]. (C) Subjective assessment of temporal salience. Average subjective aversion (1–5 scale) follows a nonlinear profile. Above 130 Hz (i.e. the discretization limit), aversive response linearly reflects stimulus frequency (red line in inset). Below this limit, aversive responses are enhanced in the roughness range and are maximal at 40 Hz (nonlinear profile, green). Error bars indicate s.e.m. * and *** indicate significant *p*-values at 0.05 and 0.001. Adapted from Arnal *et al*. [5]. (D) Sound detection in noise. Participants' performance (accuracy, reaction time and efficiency) follows an inverted U-shaped curve, optimally performing for amplitude-modulated sounds situated below the sampling limit in the roughness range. Efficiency is defined as the sum of *z*-scored accuracy and reaction speed. Quadratic fit (red parabolic curve) is calculated on efficiency data (L.H. Arnal, 2021, unpublished data).

activation of the sympathetic nervous system [102]. Recent studies on hospital alarm systems, particularly in perioperative and critical care settings, suggest that these sounds may even impair patient recovery [79,103]. Whether this effect is attributable to the aversive features of these sounds needs clarification. Although rare in natural and artificial environments, rough sounds significantly affect human behavioural, perceptual and emotional responses, but in the absence of studies specifically focusing on this feature, its effect on human health remains to be clarified.

7. Salience and exogenous attention

Much of what we perceive as humans is shaped by top–down, endogenous attention, which serves to prioritize goal-oriented sensory sampling. However, certain stimuli 'pop out' and can trigger various reactions (emotional, reflex) even when outside the focus of attention. The way the auditory system prioritizes sensitivity to certain sound frequencies and acoustic features and ensuing reactions or emotions may differ across features. Salience is a key notion in understanding perception and ensuing behavioural reactions. This essential sensory attribute primarily determines whether a stimulus will be perceived or not and whether it may capture attentional resources ahead of other inputs. Although there is a long history of investigation into exogenous attention and salience in the visual modality [104–106] and of emotional responses to vocalizations in the auditory modality [107–112]—both primarily in humans—whether specific physical features (beyond stimulus intensity) can make stimuli more or less salient in other modalities is unclear. The neural mechanisms underlying attentional capture by rough features—whether acoustic or tactile—remain understudied. Furthermore, roughness is unlikely to be the only perceptual consequence of overblown production systems, and future research should explore additional co-varying features that may contribute to perceptual salience in a similar or complementary manner.

In the following, we focus our review on recent evidence suggesting that rough sounds target and efficiently impact the cortical brain network involved in exogenous attention, the *Salience Network* (SN). We hypothesize that these circuits emerge from specific neural *salience pathways* that constitute the entry point to a general arousal system, the *salience system* (figure 3C).

Current theories posit two distinct—endogenous and exogenous—attention systems [113] that interact to guide behaviour [114]. Functional connectivity analyses of magnetic resonance imaging (fMRI) data corroborate this distinction by isolating two distinct networks in the human brain: a ventral attention SN that corresponds to exogenous attention and connects anterior cingulate and insular cortices with subcortical and limbic structures, and a dorsal endogenous attention, fronto-parietal network (FPN) linking frontal and parietal cortical regions [115]. Activity in these networks relates to distinct functions: while the FPN affects executive-task performance, the SN plays a role in monitoring relevant exogenous events. Interestingly, the SN indexes individuals' anxiety and is similarly activated by aversiveness or pain [116], supporting the view of a general, amodal *salience system* involved in modulating the perceiver's arousal [117,118] in a bottom-up, driven manner. This *salience system* may presumably be shared by vertebrates [116] and may constitute an ancestral means for initiating behavioural and arousal responses (see also [117,118]).

Whether the postulated salience system is situated downstream of canonical sensory systems or whether another *salience pathway*—bypassing the more contextual processes elaborated at the cortical level—directly activates the salience system remains uncertain [119–121]. While the existence of direct pathways has been postulated and tested in the visual [122] and auditory [108] modalities, evidence for a non-canonical pathway to the salience system remained scarce (or marginal) until recently. In the following, we will review a series of studies from animal and human experiments suggesting that neural responses to salient, rough sounds are hardly compatible with the sole recruitment of the classical auditory system. Instead, these neural patterns may suggest *a dual pathway model* in which these sounds can additionally recruit ancillary, amodal pathways particularly sensitive to the temporal features of rough sounds. In turn, given the ancestral nature of these arousal/salience systems, we propose that the preservation of this acoustic niche for alarm signalling across species may reflect the conservation of these ancient and vital circuits throughout evolution (see figure 3C).

8. Neural responses to rough sounds

Neuroimaging studies in humans have shown that rough, unpleasant sounds not only activate the auditory cortices but also the amygdala, an area involved in fear and danger processing that shares reciprocal connections with the auditory cortex [123]. A reverse correlation analysis shows that rough features efficiently activate the amygdala regardless of the context, whether vocal, artificial or music [4]. Beyond the amygdala, other regions such as the anterior insula area and the anterior cingulate cortex display greater fMRI blood oxygen level-dependent signal (i.e. higher oxygen consumption reflecting an increase in brain activity) during the perception of crying infants [83]. Unpleasant, rough sounds (nails on a chalkboard, metal scraping, dentist drill, scratchy violin) can recruit a widespread network of brain regions, including limbic and paralimbic areas such as the amygdala, the nucleus accumbens (which can mediate reward or aversion processing [124]), the insula (involved in affective/emotional processing and autonomic functions [125]), the putamen (learning, reward, cognitive processing [126]), the thalamus and the cerebellum [5,127]. Consistent with Helmholtz's original definition of roughness, dissonant sounds elicit unpleasantness, possibly via the recruitment of paralimbic and neocortical regions, including the para-hippocampus and precuneus (involved in memory, sensory processing, attention) [128]. Interestingly, human screams also have privileged access during sleep and increase slow theta waves and spindles [129], suggesting that roughness, as featured in alarm clocks and buzzers, may target sleep-related arousal systems. By analogy with the assumption that visual stimuli can impact subcortical regions such as the amygdala through direct subcortical pathways, rough sounds may be directly routed to subcortical and limbic regions involved in aversive emotional reactions, pre-empting the more detailed analyses performed by higher-level cortical networks.

9. Recruitment of a salience system

While the auditory system is the principal and decisive recipient of acoustic stimulation, the potential role of low-level or ancillary emotional pathways needs to be investigated more thoroughly. By measuring neural responses to click trains of varying rates during intracranial recordings in humans, we gathered evidence that temporally salient sounds synchronize a widespread network of subcortical nuclei (amygdala and hippocampus) and cortico-limbic regions that belong to the ventral attention SN (figure 3B). More specifically, acoustic temporal modulations between 30 and 80 Hz induce nonlinear enhancements in affective responses (subjective aversion) commensurate with increased neural synchronization. The neural pattern spans a network of temporal lobe areas, including the hippocampus, amygdala and insula, all particularly relevant to salience and affective processing. Revealing consistent nonlinear effects of sound roughness at the neural, affective and behavioural levels, this work consistently exposed a clear dichotomy in auditory and salience networks in aversive responses to this feature (see figure 3C and [5]). These findings have been replicated in several recent studies corroborating that steady-state entrainment in the roughness range (and particularly at 40 Hz) is not restricted to classical auditory regions but recruits a much larger network of medial brain regions [5,130–132]. Altogether, the widespread and sustained neural patterns in response to rough sounds evidenced in intracranial studies are incompatible with the sole recruitment of the classical auditory system. These observations, departing from typical routing and processing schemes in classical auditory pathways, suggest that these sounds additionally entrain neural activity in a *salience system* via non-canonical pathways leading to medial cortical and limbic regions [5,130] (figure 3C).

Interestingly, numerous studies in the literature demonstrate that responses to sounds in the 40 Hz range are affected in various clinical conditions (e.g. anxiety, Alzheimer's disease, schizophrenia [133]) or by pharmacological manipulations, e.g.
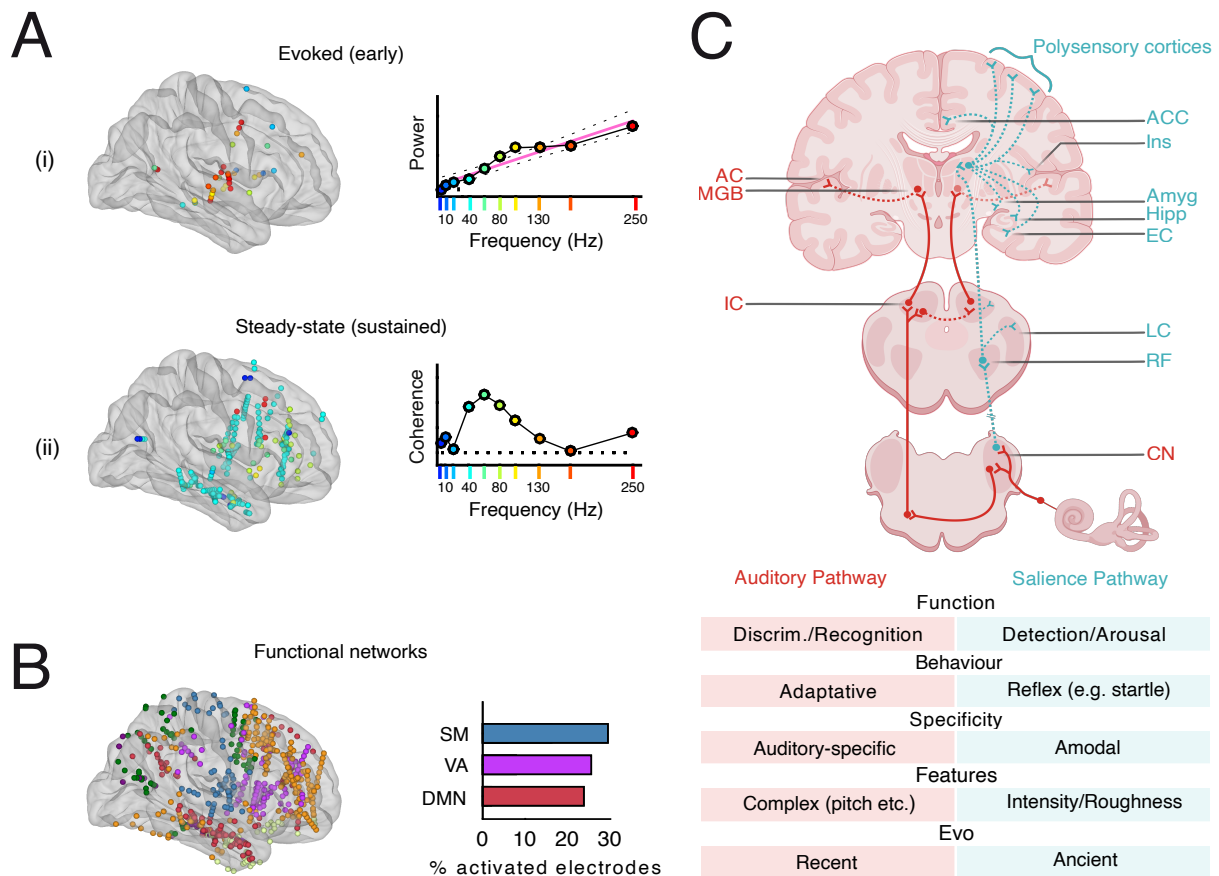
Figure 3. Dual pathways hypothesis. (A) The nonlinear patterns of subjective aversion (see figure 2) are best explained by the combination of (i) the linear recruitment of the classical auditory system as a function of click train frequency (top panels) and (ii) the nonlinear, sustained (steady-state) entrainment of neural responses in a widespread network of limbic and subcortical brain regions (bottom panels). (B) Functional networks analyses that show rough frequencies preferentially entrain neural activity in intracranial electrodes belonging to the sensorimotor network (SM), the ventral attention/salience network (VA/SN) and the default mode network (DMN), but not the dorsal attention/fronto-parietal network. Adapted from Arnal *et al*. [5]. (C) Dual neural pathways and related functions of the classical auditory system (red) and the salience system (blue). According to our hypothesis, the neural patterns generated by rough sounds are best compatible with the recruitment of a non-canonical, reticular pathway. This pathway arises in the dorsal cochlear nucleus (CN) and relays fast temporal acoustic patterns to various subcortical and cortical regions of the SN via a series of diffuse nuclei involving the reticular formation (RF). We suggest that non-canonical pathways play a crucial role in activating arousal and triggering fast responses to abrupt changes (acoustic transients) and can be probed using temporally salient, rough sounds. Table: the dual pathway hypothesis makes a series of predictions with regard to the respective functions and features of the two sensory systems. ACC, anterior cingulate cortex; Ins, insula; AC, auditory cortex; MGB, medial geniculate body; Amyg, amygdala; Hipp, hippocampus; EC, entorhinal cortex; IC, inferior colliculus; LC, locus coeruleus; Discrim., discrimination.

[134]. Of note, these diseases are rarely considered as selectively altering auditory functions but often present with less specific impairments involving attention, sleep or wakefulness—that can, but do not necessarily involve audition. Our hypothesis raises the intriguing possibility that impaired entrainment at 40 Hz in these conditions might involve defects in the salience instead of the auditory system.

10. Non-classical pathways: the dual pathway hypothesis

In the auditory domain, loud sounds induce arousing startle responses that have been known and extensively studied for a long time in humans and animals [135,136]. Recent studies in rodents showed that, beyond the fast generation of the startle motor reflex, such sounds further propagate via a non-primary auditory, 'reticular-limbic' system into deep brain regions, including the forebrain [137] and entorhinal cortex [138] (see also table figure 3C). This pattern of projections may account for the propensity of these sounds to ignite sympathetic physiological responses. Interestingly, neurons along this pathway can respond to broadband noise (but not to pure tones) at softer intensities than are required to elicit a motor startle reflex, thereby ensuring the efficient transmission of relevant signals to brain regions involved in fast (not necessarily reflex), adaptive reactions and learning [138]. This model is also supported by a series of seminal works performed in animal models in the 1960s, showing that brain and behavioural responses to sounds can be maintained despite the physical destruction of canonical auditory connections in the inferior colliculus [139]. Although more work is needed to clarify the exact pathways, it is possible that this non-canonical system corresponds to the ascending reticular activating system (ARAS), which is also characterized in human studies [140], although the terms 'non-specific' or 'extralemniscal' [141–143] may also refer to a similar notion and system. The

ARAS system is principally well known for its role in sleep regulation and arousal activation [144], evoking waking from sleep in response to sensory stimuli and enhancing arousal in response to salient stimuli (see table figure 3C and [145]). This system is considered an evolutionarily ancestral and general arousal system possibly involving neuronal reticular pathways that have been hypothesized to evolve from medullary cells of teleost fish [117,118]. Until recently, we lacked both experimental evidence and linking hypotheses to support the view that this system plays a foremost role in perception and attention. However, adding to the discovery of the SN in humans [146,147], a series of very recent studies in animal models have revealed that the ARAS projects to various limbic regions and dynamically regulates the SN [148]. These crucial observations provide timely anatomical and functional evidence in favour of a salience pathway originating in the ARAS and ultimately projecting to (and exogenously controlling) the SN.

While recent research has elucidated the perceptual and neural processing of roughness in humans, there is a notable lack of comparable studies in nonhuman animals. Whether these sounds are perceived and processed by the proposed salience system in other species remains speculative at this point. Future comparative research should aim to systematically relate the use of roughness in alarm vocal signals to their behavioural, affective and neural effects across a broader range of species, e.g. from mammals to reptiles, to better understand the evolutionary and cross-species relevance of these features for alarm communication.

11. Conclusions

Several neural mechanisms may account for the propensity of rough and nonlinear vocalizations to enhance arousal in the receiver. Stemming from the observation that nonlinear vocal phenomena in mammals exhibit chaotic features, several authors have argued that the resulting sounds enhance attentional capture and prevent habituation owing to their unpredictability [8,40]. While this hypothesis is interesting and possibly correct, the evidence collected in this review points to another—although compatible—sensory phenomenon, pertaining to the capacity of these rapid temporal features, such as repetitions of acoustic transients or amplitude modulations, to entrain neural responses in a sustained manner over time. In this view, temporal salience, akin to a strobe light or tactile vibration effect, arises from the bombardment of sensory systems just below their sampling (fusion) limit, resulting in widespread synchronization throughout regions involved in salience processing. According to the dual pathways hypothesis, transient acoustic stimuli are not only carried by the auditory system but also via an evolutionarily ancestral, amodal salience system that is responsible for modulating the receiver's arousal state (figure 3C). That fast repetitive features can be found in the communication signals of a wide array of vertebrate species, such as fishes and toads (figure 1), suggests that they may have been conserved over the course of evolution, which could date back to early adaptations of primordial sensory pathways [7]. This view contrasts with less urgent, more complex—e.g. tonal or harmonic—signals that arguably arose later to populate other acoustic niches of the audible spectrum and require more advanced processing involving auditory cortical regions.

In our hypothesis, the salience system is specifically sensitive to a restricted range of stimulation rates and may be probed using temporally salient sounds. If this system is mostly sensitive to impulses, by repeating impulses in time at various frequencies, it should be possible to identify the preferred resonance rate of this system. Therefore, temporally 'enriching' stimuli up to the sampling limit of the salience system might enhance stimulus detectability and increase subjective salience. Beyond this limit, events are not perceived as discrete but rather as continuous (pitch) and may not be perceived as salient any more. Altogether, these findings point to a supramodal neural correlate of temporal salience: the aversive sensation induced by rough sounds results from the persistent, exogenous synchronization of large-scale networks involved in salience—rather than specifically auditory—processing. This also suggests that the negative perception induced by rough sounds such as dissonant intervals [149], alarm sounds [4] or annoying vocal effects (e.g. vocal fry [150]) might result from their capacity to exogenously hijack brain networks involved in salience, aversion or pain processing [116]. In addition to the large extent of spatial synchronization patterns, it is remarkable that these networks preferentially resonate in a frequency range (30–150 Hz) that matches a well-known endogenous brain rhythm, the so-called gamma band [151]. Neuronal synchronization in the gamma range has been proposed as a mechanism to subserve the selective routing of bottom-up information in the brain [152,153]. Consequently, stimulus-driven, privileged entrainment in the gamma range by rough sounds may reflect the exogenous recruitment of attentional and arousal-related brain regions to ultimately enhance perceptual salience. Another interesting feature of this phenomenon is that, unlike early evoked auditory responses, steady-state sustained electrophysiological responses in the gamma range do not seem to habituate: their magnitude remains stable even after several repetitions of a sound (L.H. Arnal, 2021, unpublished result). This aspect is particularly relevant in the context of vocal communication because preventing neural habituation may ensure that the processing of alarm sounds is always prioritized over other environmental events.

This theoretical framework also has interesting implications for the understanding of salience in acoustic communication and supports the hypothesis that temporally enriching sounds in time—in the roughness range—amplify sensory salience and improve neural and behavioural efficiency. This nicely accounts for the observed convergence of roughness as a privileged acoustic niche to warn conspecifics [4]. Showing that such sounds recruit salience systems in the human brain and enhance perception, we confirm the fitness of these sounds to ultimately promote the efficient transmission of signals that aim at grabbing receivers' attention. This further provides evidence in favour of the hypothesis that the use of roughness in alarm signalling is not an epiphenomenon of vocal production. Instead, the selection of communicative features depends on their propensity to enhance the detectability of vocal utterances and to induce adapted behavioural responses. In this view, the use of roughness in alarm signals may reflect an adaptation of communication to the receiver's auditory sampling constraints to hijack

her brain, enhance her perception of incoming danger and manipulate her reactions from a distance to ultimately promote survival.

References

1. Briefer EF. 2012 Vocal expression of emotions in mammals: mechanisms of production and evidence. *J. Zool*. **288**, 1–20. (doi:10.1111/j.1469-7998.2012.00920.x)

2. Mendl M, Burman O, Paul E. 2010 An integrative and functional framework for the study of animal emotion and mood. *Proc. R. Soc. B* **277**, 2895–2904. (doi:10.1098/rspb.2010.0303)

3. Pisanski K, Reby D, Oleszkiewicz A. 2024 Humans need auditory experience to produce typical volitional nonverbal vocalizations. *Commun. Psychol* **2**, 1–15. (doi:10.1038/s44271-024-00104-6)

4. Arnal LH, Flinker A, Kleinschmidt A, Giraud AL, Poeppel D. 2015 Human screams occupy a privileged niche in the communication soundscape. *Curr. Biol*. **25**, 2051–2056. (doi:10.1016/j.cub.2015.06.043)

5. Arnal LH, Kleinschmidt A, Spinelli L, Giraud AL, Mégevand P. 2019 The rough sound of salience enhances aversion through neural synchronisation. *Nat. Commun*. **10**, 3671. (doi:10.1038/s41467-019-11626-7)

6. Lieberman P. 1985 The physiology of cry and speech in relation to linguistic behavior. In *Infant crying: theoretical and research perspectives* (eds BM Lester, CF Zachariah Boukydis), pp. 29–57. Boston, MA: Springer US. (doi:10.1007/978-1-4613-2381-5_2)

7. Owren MJ, Rendall D. 2001 Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evol. Anthropol. Issues News Rev*. **10**, 58–71. (doi:10.1002/evan.1014)

8. Fitch WT, Neubauer J, Herzel H. 2002 Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Anim. Behav*. **63**, 407–418. (doi:10.1006/anbe.2001.1912)

9. Lingle S, Wyman MT, Kotrba R, Teichroeb LJ, Romanow CA. 2012 What makes a cry a cry? A review of infant distress vocalizations. *Curr. Zool*. **58**, 698–726. (doi:10.1093/czoolo/58.5.698)

10. Cornec C, Mathevon N, Pisanski K, Entani D, Monghiemo C, Bola B, Planas-Bielsa V, Reby D, Levréro F. 2024 Human infant cries communicate distress and elicit sex stereotypes: cross cultural evidence. *Evol. Hum. Behav*. **45**, 48–57. (doi:10.1016/j.evolhumbehav.2023.08.004)

11. Zeskind PS, Lester BM. 1978 Acoustic features and auditory perceptions of the cries of newborns with prenatal and perinatal complications. *Child Dev*. **49**, 580–589. (doi:10.2307/1128224)

12. Corvin S, Massenet M, Hardy A, Patural H, Peyron R, Fauchon C, Mathevon N. 2025 Nonlinear acoustic phenomena affect the perception of pain in human baby cries. *Phil. Trans. R. Soc. B* **380**, 20240023. (doi:10.1098/rstb.2024.0023)

13. Koutseff A, Reby D, Martin O, Levrero F, Patural H, Mathevon N. 2017 The acoustic space of pain: cries as indicators of distress recovering dynamics in pre-verbal infants. *Bioacoustics* **27**, 1–13. (doi:10.1080/09524622.2017.1344931)

14. Van Lierde KM, Dijckmans J, Scheffel L, Behlau M. 2012 Type and severity of pain during phonation in professional voice users and nonvocal professionals. *J. Voice* **26**, 671.(doi:10.1016/j.jvoice.2011.11.008)

15. Bradley MM, Lang PJ. 2000 Affective reactions to acoustic stimuli. *Psychophysiology* **37**, 204–215. (doi:10.1111/1469-8986.3720204)

16. Schwartz JW, Gouzoules H. 2019 Decoding human screams: perception of emotional arousal from pitch and duration. *Behaviour* **156**, 1283–1307. (doi:10.1163/1568539x-00003566)

17. Trevor C, Arnal LH, Frühholz S. 2020 Terrifying film music mimics alarming acoustic feature of human screams. *J. Acoust. Soc. Am*. **147**, EL540–EL545. (doi:10.1121/10.0001459)

18. Beaurenaut M, Tokarski E, Dezecache G, Grèzes J. 2020 The 'Threat of Scream' paradigm: a tool for studying sustained physiological and subjective anxiety. *Sci. Rep*. **10**, 12496. (doi:10.1038/s41598-020-68889-0)

19. Frühholz S, Dietziker J, Staib M, Trost W. 2021 Neurocognitive processing efficiency for discriminating human non-alarm rather than alarm scream calls. *PLoS Biol*. **19**, e3000751. (doi:10.1371/journal.pbio.3000751)

20. Seyfarth RM, Cheney DL, Marler P. 1980 Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Anim. Behav*. **28**, 1070–1094. (doi:10.1016/s0003-3472(80)80097-2)

21. Anikin A, Pisanski K, Reby D. 2020 Do nonlinear vocal phenomena signal negative valence or high emotion intensity? *R. Soc. Open Sci*. **7**, 201306. (doi:10.1098/rsos.201306)

22. Titze I. 2000 *Principles of voice production,* second printing, pp. 245–251. Iowa City, IA: National Center for Voice and Speech.

23. Fant G. 1971 *AcousticTheory of Speech Production: With Calculations Based on X-Ray Studies of Russian Articulations (no. 2)*. Berlin, Germany: Walter de Gruyter. (doi:10.1515/9783110873429)

24. Rendall D, Notman H, Owren MJ. 2009 Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *J. Acoust. Soc. Am*. **125**, 1792–1805. (doi:10.1121/1.3068453)

25. Stoeger AS, Charlton BD, Kratochvil H, Fitch WT. 2011 Vocal cues indicate level of arousal in infant African elephant roars. *J. Acoust. Soc. Am*. **130**, 1700–1710. (doi:10.1121/1.3605538)

26. Massenet M, Anikin A, Pisanski K, Reynaud K, Mathevon N, Reby D. 2022 Nonlinear vocal phenomena affect human perceptions of distress, size and dominance in puppy whines. *Proc. R. Soc. B* **289**, 20220429. (doi:10.1098/rspb.2022.0429)

27. Townsend SW, Manser MB. 2010 The function of nonlinear phenomena in meerkat alarm calls. *Biol. Lett.* **7**, 47–49. (doi:10.1098/rsbl.2010.0537)

28. Gouzoules S, Gouzoules H, Marler P. 1984 Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Anim. Behav.* **32**, 182–193. (doi:10.1016/s0003-3472(84)80336-x)

29. McRae TR. 2020 A review of squirrel alarm-calling behavior: What we know and what we do not know about how predator attributes affect alarm calls. *Anim. Behav. Cogn.* **7**, 168–191. (doi:10.26451/abc.07.02.11.2020)

30. Karp D, Manser MB, Wiley EM, Townsend SW. 2013 Nonlinearities in meerkat alarm calls prevent receivers from habituating. *Ethology* **120**, 189–196. (doi:10.1111/eth.12195)

31. Macedonia JM. 1993 The vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatol. Int. J. Primatol.* **61**, 186–217. (doi:10.1159/000156749)

32. Blumstein DT, Richardson DT, Cooley L, Winternitz J, Daniel JC. 2008 The structure, meaning and function of yellow-bellied marmot pup screams. *Anim. Behav.* **76**, 1055–1064. (doi:10.1016/j.anbehav.2008.06.002)

33. Tokuda I, Riede T, Neubauer J, Owren MJ, Herzel H. 2024 Nonlinear analysis of irregular animal vocalizations. *J. Acoust. Soc. Am.* **111**, 2908–2919. (doi:10.1121/1.1474440)

34. Massenet M, Pisanski K, Reynaud K, Mathevon N, Reby D, Anikin A. 2025 Acoustic context and dynamics of nonlinear phenomena in mammalian calls: the case of puppy whines. *Phil. Trans. R. Soc. B* **380**, 20240022. (doi:10.1098/rstb.2024.0022)

35. Mende W, Herzel H, Wermke K. 1990 Bifurcations and chaos in newborn infant cries. *Phys. Lett.* A **145**, 418–424. (doi:10.1016/0375-9601(90)90305-8)

36. Facchini A, Bellieni CV, Marchettini N, Pulselli FM, Tiezzi EBP. 2005 Relating pain intensity of newborns to onset of nonlinear phenomena in cry recordings. *Phys. Lett.* A **338**, 332–337. (doi:10.1016/j.physleta.2005.02.048)

37. Pisanski K, Bryant GA, Cornec C, Anikin A, Reby D. 2022 Form follows function in human nonverbal vocalisations. *Ethol. Ecol. Evol.* **34**, 303–321. (doi:10.1080/03949370.2022.2026482)

38. Lingle S, Riede T. 2014 Deer mothers are sensitive to infant distress vocalizations of diverse mammalian species. *Am. Nat.* **184**, 510–522. (doi:10.1086/677677)

39. Anikin A. 2020 The perceptual effects of manipulating nonlinear phenomena in synthetic nonverbal vocalizations. *Bioacoustics* **29**, 226–247. (doi:10.1080/09524622.2019.1581839)

40. Blumstein DT, Récapet C. 2009 The sound of arousal: the addition of novel non-linearities increases responsiveness in marmot alarm calls. *Ethology* **115**, 1074–1081. (doi:10.1111/j.1439-0310.2009.01691.x)

41. Blumstein DT. 2025 Nonlinear phenomena in marmot alarm calls: a mechanism encoding fear? *Phil. Trans. R. Soc. B* **380**, 20240008. (doi:10.1098/rstb.2024.0008)

42. Fuchs E, Beeck VC, Baotic A, Stoeger AS. 2021 Acoustic structure and information content of trumpets in female Asian elephants (*Elephas maximus*). *PLoS One* **16**, e0260284. (doi:10.1371/journal.pone.0260284)

43. Marx A, Lenkei R, Pérez Fraga P, Bakos V, Kubinyi E, Faragó T. 2021 Occurrences of non-linear phenomena and vocal harshness in dog whines as indicators of stress and ageing. *Sci. Rep.* **11**, 4468. (doi:10.1038/s41598-021-83614-1)

44. Riede T, Herzel H, Mehwald D, Seidner W, Trumler E, Böhme G, Tembrock G. 2000 Nonlinear phenomena in the natural howling of a dog–wolf mix. *J. Acoust. Soc. Am.* **108**, 1435–1442. (doi:10.1121/1.1289208)

45. Townsend SW, Manser MB. 2011 The function of nonlinear phenomena in meerkat alarm calls. *Biol. Lett.* **7**, 47–49. (doi:10.1098/rsbl.2010.0537)

46. Mann DA, O'Shea TJ, Nowacek DP. 2006 Nonlinear dynamics in manatee vocalizations. *Mar. Mammal Sci.* **22**, 548–555. (doi:10.1111/j.1748-7692.2006.00036.x)

47. Riede T, Arcadi AC, Owren MJ. 2007 Nonlinear acoustics in the pant hoots of common chimpanzees (*Pan troglodytes*): vocalizing at the edge. *J. Acoust. Soc. Am.* **121**, 1758–1767. (doi:10.1121/1.2427115)

48. Herbst CT, Tokuda IT, Nishimura T, Ternström S, Ossio V, Levy M, Fitch T, Dunn JC. 2025 Monkey yodels'—frequency jumps in new world monkey vocalizations greatly surpass human vocal register transitions. *Phil. Trans. R. Soc. B* **380**, 20240005. (doi:10.1098/rstb.2024.0005)

49. Charlton BD. 2015 The acoustic structure and information content of female koala vocal signals. *PLoS One* **10**, e0138670. (doi:10.1371/journal.pone.0138670)

50. Stoeger AS, Baotic A, Li D, Charlton BD. 2012 Acoustic features indicate arousal in infant giant panda vocalisations. *Ethology* **118**, 896–905. (doi:10.1111/j.1439-0310.2012.02080.x)

51. Volodin I, Matrosova V, Frey R, Kozhevnikova J, Isaeva I, Volodina E. 2018 Altai pika (*Ochotona alpina*) alarm calls: individual acoustic variation and the phenomenon of call-synchronous ear folding behavior. *Naturwissenschaften* **105**, 40. (doi:10.1007/s00114-018-1567-8)

52. Volodin IA, Volodina EV, Frey R, Gogoleva SS, Palko IV, Rozhnov VV. 2017 Acoustic structure of alarm calls in Indian sambar (*Rusa unicolor*) and Indian muntjac (*Muntiacus vaginalis*) in South Vietnam. *Dokl. Biol. Sci. Proc. Acad. Sci. USSR Biol. Sci. Sect.* **474**, 110–113. (doi:10.1134/S0012496617030061)

53. Reby D, Wyman MT, Frey R, Passilongo D, Gilbert J, Locatelli Y, Charlton BD. 2016 Evidence of biphonation and source-filter interactions in the bugles of male North American wapiti (*Cervus canadensis*). *J. Exp. Biol.* **219**, 1224–1236. (doi:10.1242/jeb.131219)

54. Green AC, Clark CEF, Lomax S, Favaro L, Reby D. 2020 Context-related variation in the peripartum vocalisations and phonatory behaviours of Holstein-Friesian dairy cows. *Appl. Anim. Behav. Sci.* **231**, 105089. (doi:10.1016/j.applanim.2020.105089)

55. Schneider JN, Anderson RE. 2011 Tonal vocalizations in the red wolf (*Canis rufus*): Potential functions of nonlinear sound production. *J. Acoust. Soc. Am.* **130**, 2275–2284. (doi:10.1121/1.3628331)

56. Tyson RB, Nowacek DP, Miller PJO. 2007 Nonlinear phenomena in the vocalizations of North Atlantic right whales (*Eubalaena glacialis*) and killer whales (*Orcinus orca*). *J. Acoust. Soc. Am.* **122**, 1365–1373. (doi:10.1121/1.2756263)

57. Cazau D, Adam O, Aubin T, Laitman JT, Reidenberg JS. 2016 A study of vocal nonlinearities in humpback whale songs: from production mechanisms to acoustic analysis. *Sci. Rep.* **6**, 31660. (doi:10.1038/srep31660)

58. Samarra FIP. 2015 Variations in killer whale food-associated calls produced during different prey behavioural contexts. *Behav. Process.* **116**, 33–42. (doi:10.1016/j.beproc.2015.04.013)

59. Figueiredo LD de, Maciel I, Viola FM, Savi MA, Simão SM. 2023 Nonlinear features in whistles produced by the short-beaked common dolphin (*Delphinus delphis*) off southeastern Brazil. *J. Acoust. Soc. Am.* **153**, 2436. (doi:10.1121/10.0017883)

60. Blesdoe EK, Blumstein DT. 2014 What is the sound of fear? Behavioral responses of white-crowned sparrows *Zonotrichia leucophrys* to synthesized nonlinear acoustic phenomena. *Curr. Zool.* **60**, 534–541. (doi:10.1093/czoolo/60.4.534)

61. Amador A, Mindlin GB, Elemans CPH. 2025 Transitions and tricks: nonlinear phenomena in the avian voice. *Phil. Trans. R. Soc. B* **380**, 20240007. (doi:10.1098/rstb.2024.0007)

62. Zhang F, Zhao J, Feng AS. 2017 Vocalizations of female frogs contain nonlinear characteristics and individual signatures. *PLoS One* **12**, e0174815. (doi:10.1371/journal.pone.0174815)

63. Kamiloğlu RG, Slocombe KE, Haun DBM, Sauter DA. 2020 Human listeners' perception of behavioural context and core affect dimensions in chimpanzee vocalizations. *Proc. R. Soc. B* **287**, 20201148. (doi:10.1098/rspb.2020.1148)

64. Debracque C, Clay Z, Grandjean D, Gruber T. 2023 Humans recognize affective cues in primate vocalizations: Acoustic and phylogenetic perspectives. *Sci. Rep* **13**, 10900. (doi:10.1038/s41598-023-37558-3)

65. Schwartz JW, Gouzoules H. 2022 Humans read emotional arousal in monkey vocalizations: evidence for evolutionary continuities in communication. *PeerJ* **10**, e14471. (doi:10.7717/peerj.14471)

66. Filippi P *et al*. 2017 Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: evidence for acoustic universals. *Proc. R. Soc. B* **284**, 20170990. (doi:10.1098/rspb.2017.0990)

67. Thévenet J, Papet L, Coureaud G, Boyer N, Levréro F, Grimault N, Mathevon N. 2023 Crocodile perception of distress in hominid baby cries. *Proc. R. Soc. B* **290**, 20230201. (doi:10.1098/rspb.2023.0201)

68. Singh NC, Theunissen FE. 2003 Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am*. **114**, 3394–3411. (doi:10.1121/1.1624067)

69. Elliott TM, Theunissen FE. 2009 The modulation transfer function for speech intelligibility. *PLoS Comput. Biol*. **5**, e1000302. (doi:10.1371/journal.pcbi.1000302)

70. Theunissen FE, Elie JE. 2014 Neural processing of natural sounds. *Nat. Rev. Neurosci*. **15**, 355–366. (doi:10.1038/nrn3731)

71. Chi T, Ru P, Shamma SA. 2005 Multiresolution spectrotemporal analysis of complex sounds. *J. Acoust. Soc. Am*. **118**, 887–906. (doi:10.1121/1.1945807)

72. Taffou M, Suied C, Viaud-Delmon I. 2021 Auditory roughness elicits defense reactions. *Sci. Rep*. **11**, 956. (doi:10.1038/s41598-020-79767-0)

73. Terhardt E. 1978 Psychoacoustic evaluation of musical sounds. *Percept. Psychophys*. **23**, 483–492. (doi:10.3758/bf03199523)

74. Pressnitzer D, McAdams S, Winsberg S, Fineberg J. 2000 Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Percept. Psychophys*. **62**, 66–80. (doi:10.3758/bf03212061)

75. Vassilakis P. 2005 Auditory roughness as a means of musical expression. *Sel. Rep. Ethnomusicol*. **7**, 119.

76. Terhardt E. 1974 On the perception of periodic sound fluctuations (Roughness). *Acta Acust. United Acust*. **30**, 201–213.

77. Blumstein DT, Davitian R, Kaye PD. 2010 Do film soundtracks contain nonlinear analogues to influence emotion? *Biol. Lett*. **6**, 751–754. (doi:10.1098/rsbl.2010.0333)

78. McDermott JH, Schultz AF, Undurraga EA, Godoy RA. 2016 Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* **535**, 547–550. (doi:10.1038/nature18635)

79. Watson N, Shah A, Patel P, Soni S. 2023 Alarming tones: the unexplored phenomenon of auditory roughness. *Br. J. Anaesth*. **131**, e103–e106. (doi:10.1016/j.bja.2023.05.010)

80. Vassilakis PN, Kendall RA. 2010 Psychoacoustic and cognitive aspects of auditory roughness: definitions, models, and applications. In *Proceedings Volume 7527, Human Vision and Electronic Imaging XV* (San Jose, CA, 17 February 2010), pp. 203–209. SPIE. (doi:10.1117/12.845457)

81. Di Stefano N, Spence C. 2022 Roughness perception: a multisensory/crossmodal perspective. *Atten. Percept. Psychophys*. **84**, 2087–2114. (doi:10.3758/s13414-022-02550-y)

82. Zhao S, Yum NW, Benjamin L, Benhamou E, Yoneya M, Furukawa S, Dick F, Slaney M, Chait M. 2019 Rapid ocular responses are modulated by bottom-up-driven auditory salience. *J. Neurosci*. **39**, 7703–7714. (doi:10.1523/jneurosci.0776-19.2019)

83. Li T, Horta M, Mascaro JS, Bijanki K, Arnal LH, Adams M, Barr RG, Rilling JK. 2018 Explaining individual variation in paternal brain responses to infant cries. *Physiol. Behav*. **193**, 43–54. (doi:10.1016/j.physbeh.2017.12.033)

84. Kurth E, Kennedy HP, Spichiger E, Hösli I, Zemp Stutz E. 2011 Crying babies, tired mothers: What do we know? A systematic review. *Midwifery* **27**, 187–194. (doi:10.1016/j.midw.2009.05.012)

85. Möller EL, de Vente W, Rodenburg R. 2019 Infant crying and the calming response: Parental versus mechanical soothing using swaddling, sound, and movement. *PLoS One* **14**, e0214548. (doi:10.1371/journal.pone.0214548)

86. Hu L, Han B, Zhao X, Mi L, Song Q, Wang J, Song T, Huang C. 2016 Chronic early postnatal scream sound stress induces learning deficits and NMDA receptor changes in the hippocampus of adult mice. *NeuroReport* **27**, 397–403. (doi:10.1097/wnr.0000000000000552)

87. Hu L, Zhao X, Yang J, Wang L, Yang Y, Song T, Huang C. 2014 Chronic scream sound exposure alters memory and monoamine levels in female rat brain. *Physiol. Behav*. **137**, 53–59. (doi:10.1016/j.physbeh.2014.06.012)

88. Gonzalez-Palomares E, Boulanger-Bertolus J, Dupin M, Mouly AM, Hechavarria JC. 2023 Amplitude modulation pattern of rat distress vocalisations during fear conditioning. *Sci. Rep*. **13**, 11173. (doi:10.1038/s41598-023-38051-7)

89. Hechavarría JC, Jerome Beetz M, García-Rosales F, Kössl M. 2020 Bats distress vocalizations carry fast amplitude modulations that could represent an acoustic correlate of roughness. *Sci. Rep*. **10**, 7332. (doi:10.1038/s41598-020-64323-7)

90. Bastian A, Schmidt S. 2008 Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J. Acoust. Soc. Am*. **124**, 598–608. (doi:10.1121/1.2924123)

91. Carter G, Schoeppler D, Manthey M, Knörnschild M, Denzinger A. 2015 Distress calls of a fast-flying bat (*Molossus molossus*) provoke inspection flights but not cooperative mobbing. *PLoS One* **10**, e0136146. (doi:10.1371/journal.pone.0136146)

92. Taylor D, Gustafsson E, Dezecache G, Davila-Ross M. 2023 Directedness and engagement in chimpanzee vocal ontogeny. *Dev. Sci*. **26**, e13334. (doi:10.1111/desc.13334)

93. Postal O, Dupont T, Bakay W, Dominique N, Petit C, Michalski N, Gourévitch B. 2020 Spontaneous mouse behavior in presence of dissonance and acoustic roughness. *Front. Behav. Neurosci*. **14**, 588834. (doi:10.3389/fnbeh.2020.588834)

94. Critchley HD, Harrison NA. 2013 Visceral influences on brain and behavior. *Neuron* **77**, 624–638. (doi:10.1016/j.neuron.2013.02.008)

95. Barrett LF, Simmons WK. 2015 Interoceptive predictions in the brain. *Nat. Rev. Neurosci*. **16**, 419–429. (doi:10.1038/nrn3950)

96. Chen WG *et al*. 2021 The emerging science of interoception: sensing, integrating, interpreting, and regulating signals within the self. *Trends Neurosci*. **44**, 3–16. (doi:10.1016/j.tins.2020.10.007)

97. Aureli F, Preston SD, de Waal FBM. 1999 Heart rate responses to social interactions in free-moving rhesus macaques (*Macaca mulatta*): a pilot study. *J. Comp. Psychol*. **113**, 59–65. (doi:10.1037//0735-7036.113.1.59)

98. Hsueh B *et al*. 2023 Cardiogenic control of affective behavioural state. *Nature* **615**, 292–299. (doi:10.1038/s41586-023-05748-8)

99. Ma J, Kanwal JS. 2014 Stimulation of the basal and central amygdala in the mustached bat triggers echolocation and agonistic vocalizations within multimodal output. *Front. Physiol*. **5**, 55. (doi:10.3389/fphys.2014.00055)

100. LeDoux JE, Farb C, Ruggiero DA. 1990 Topographic organization of neurons in the acoustic thalamus that project to the amygdala. *J. Neurosci. Off. J. Soc. Neurosci*. **10**, 1043–1054. (doi:10.1523/JNEUROSCI.10-04-01043.1990)

101. Kumar S, von Kriegstein K, Friston K, Griffiths TD. 2012 Features versus feelings: Dissociable representations of the acoustic features and valence of aversive sounds. *J. Neurosci*. **32**, 14184–14192. (doi:10.1523/jneurosci.1759-12.2012)

102. Joosen KJ, Mesman J, Bakermans-Kranenburg MJ, Pieper S, Zeskind PS, van IJzendoorn MH. 2013 Physiological reactivity to infant crying and observed maternal sensitivity. *Infancy* **18**, 414–431. (doi:10.1111/j.1532-7078.2012.00122.x)

103. Burdick KJ, Yang S, Lopez AE, Wessel C, Schutz M, Schlesinger JJ. 2023 Auditory roughness: a delicate balance. *Br. J. Anaesth*. **131**, 649–652. (doi:10.1016/j.bja.2023.07.003)

104. Chun MM, Marois R. 2002 The dark side of visual attention. *Curr. Opin. Neurobiol*. **12**, 184–189. (doi:10.1016/s0959-4388(02)00309-4)

105. Hickey C, McDonald JJ, Theeuwes J. 2006 Electrophysiological evidence of the capture of visual attention. *J. Cogn. Neurosci*. **18**, 604–613. (doi:10.1162/jocn.2006.18.4.604)

106. Hopfinger JB, West VM. 2006 Interactions between endogenous and exogenous attention on cortical visual processing. *NeuroImage* **31**, 774–789. (doi:10.1016/j.neuroimage.2005.12.049)

107. Fecteau S, Belin P, Joanette Y, Armony JL. 2007 Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage* **36**, 480–487. (doi:10.1016/j.neuroimage.2007.02.043)

108. Frühholz S, Grandjean D. 2013 Amygdala subregions differentially respond and rapidly adapt to threatening voices. *Cortex* **49**, 1394–1403. (doi:10.1016/j.cortex.2012.08.003)

109. Grandjean D, Sander D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, Vuilleumier P. 2005 The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci*. **8**, 145–146. (doi:10.1038/nn1392)

110. Morris JS, Scott SK, Dolan RJ. 1999 Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia* **37**, 1155–1163. (doi:10.1016/s0028-3932(99)00015-9)

111. Schirmer A, Kotz SA. 2006 Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci*. **10**, 24–30. (doi:10.1016/j.tics.2005.11.009)

112. Scott SK, Sauter D, McGettigan C. 2010 Brain mechanisms for processing perceived emotional vocalizations in humans. In *Handbook of behavioral neuroscience,* vol. 19 (ed. SM Brudzynski), pp. 187–197. (doi:10.1016/B978-0-12-374593-4.00019-X)

113. Posner MI, Petersen SE. 1990 The attention system of the human brain. *Annu. Rev. Neurosci*. **13**, 25–42. (doi:10.1146/annurev.ne.13.030190.000325)

114. Corbetta M, Shulman GL. 2002 Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci*. **3**, 201–215. (doi:10.1038/nrn755)

115. Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007 Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci*. **27**, 2349–2356. (doi:10.1523/jneurosci.5587-06.2007)

116. Hayes DJ, Northoff G. 2012 Common brain activations for painful and non-painful aversive stimuli. *BMC Neurosci*. **13**, 60. (doi:10.1186/1471-2202-13-60)

117. Pfaff DW, Martin EM, Faber D. 2012 Origins of arousal: roles for medullary reticular neurons. *Trends Neurosci*. **35**, 468–476. (doi:10.1016/j.tins.2012.04.008)

118. Quinkert AW, Vimal V, Weil ZM, Reeke GN, Schiff ND, Banavar JR, Pfaff DW. 2011 Quantitative descriptions of generalized arousal, an elementary function of the vertebrate brain. *Proc. Natl Acad. Sci. USA* **108 Suppl 3**, 15617–15623. (doi:10.1073/pnas.1101894108)

119. Bordi F, LeDoux JE. 1994 Response properties of single units in areas of rat auditory thalamus that project to the amygdala. I. Acoustic discharge patterns and frequency receptive fields. *Exp. Brain Res*. **98**, 261–274. (doi:10.1007/BF00228414)

120. Phelps EA, LeDoux JE. 2005 Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* **48**, 175–187. (doi:10.1016/j.neuron.2005.09.025)

121. Ledoux JE. 2000 Emotion circuits in the brain. *Ann. Rev. Neurosci*. **23**, 155–184. (doi:10.1146/annurev.neuro.23.1.155)

122. Méndez-Bértolo C, Moratti S, Toledano R, Lopez-Sosa F, Martínez-Alvarez R, Mah YH, Vuilleumier P, Gil-Nagel A, Strange BA. 2016 A fast pathway for fear in human amygdala. *Nat. Neurosci*. **19**, 1041–1049. (doi:10.1038/nn.4324)

123. Belin P, Zatorre RJ. 2015 Neurobiology: Sounding the alarm. *Curr. Biol*. **25**, R805–R806. (doi:10.1016/j.cub.2015.07.027)

124. Zhou K, Xu H, Lu S, Jiang S, Hou G, Deng X, He M, Zhu Y. 2022 Reward and aversion processing by input-defined parallel nucleus accumbens circuits in mice. *Nat. Commun*. **13**, 6244. (doi:10.1038/s41467-022-33843-3)

125. Gogolla N. 2017 The insular cortex. *Curr. Biol*. **27**, R580–R586. (doi:10.1016/j.cub.2017.05.010)

126. Ghandili M, Munakomi S. 2024 *Neuroanatomy, Putamen*. Treasure Island, FL: StatPearls Publishing.

127. Zald DH, Pardo JV. 2002 The neural correlates of aversive auditory stimulation. *NeuroImage* **16**, 746–753. (doi:10.1006/nimg.2002.1115)

128. Blood AJ, Zatorre RJ, Bermudez P, Evans AC. 1999 Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci*. **2**, 382–387. (doi:10.1038/7299)

129. Legendre G, Moyne M, Dominguez-Borras J, Kumar S, Sterpenich V, Schwartz S, Arnal LH. 2022 Scream's roughness confers a privileged access to the brain during sleep. *bioRxiv* 2022.09.05.506631. (doi:10.1101/2022.09.05.506631)

130. Farahani ED, Goossens T, Wouters J, van Wieringen A. 2017 Spatiotemporal reconstruction of auditory steady-state responses to acoustic amplitude modulations: Potential sources beyond the auditory pathway. *NeuroImage* **148**, 240–253. (doi:10.1016/j.neuroimage.2017.01.032)

131. Tada M *et al*. 2021 Global and parallel cortical processing based on auditory gamma oscillatory responses in humans. *Cereb. Cortex* **31**, 4518–4532. (doi:10.1093/cercor/bhab103)

132. Blanpain LT, Cole ER, Chen E, Park JK, Walelign MY, Gross RE, Cabaniss BT, Willie JT, Singer AC. 2024 Multisensory flicker modulates widespread brain networks and reduces interictal epileptiform discharges. *Nat. Commun*. **15**, 3156. (doi:10.1038/s41467-024-47263-y)

133. Thun H, Recasens M, Uhlhaas P. 2016 The 40-Hz auditory steady-state response in patients with schizophrenia a meta-analysis. *JAMA Psychiatry* **73**, 1145–1153. (doi:10.1001/jamapsychiatry.2016.2619)

134. Tada M *et al*. 2019 Gamma-band auditory steady-state response as a neurophysiological marker for excitation and inhibition balance: A review for understanding schizophrenia and other neuropsychiatric disorders. *Clin. EEG Neurosci*. **51**, 234–243. (doi:10.1177/1550059419868872)

135. Geyer MA, Swerdlow NR. 2001 Measurement of startle response, prepulse inhibition, and habituation. *Curr. Protoc. Neurosci*. **3**, 8. (doi:10.1002/0471142301.ns0807s03)

136. Swerdlow N, Geyer M, Braff D. 2001 Neural circuit regulation of prepulse inhibition of startle in the rat: current knowledge and future challenges. *Psychopharmacology* **156**, 194–215. (doi:10.1007/s002130100799)

137. Kim T *et al*. 2015 Cortically projecting basal forebrain parvalbumin neurons regulate cortical gamma band oscillations. *Proc. Natl Acad. Sci. USA* **112**, 3535–3540. (doi:10.1073/pnas.1413625112)

138. Zhang GW, Sun WJ, Zingg B, Shen L, He J, Xiong Y, Tao HW, Zhang LI. 2018 A non-canonical reticular-limbic central auditory pathway via medial septum contributes to fear conditioning. *Neuron* **97**, 406–417. (doi:10.1016/j.neuron.2017.12.010)

139. Galambos R, Myers RE, Sheatz GC. 1961 Extralemniscal activation of auditory cortex in cats. *Am. J. Physiol*. **200**, 23–28. (doi:10.1152/ajplegacy.1961.200.1.23)

140. Moruzzi G, Magoun HW. 1949 Brain stem reticular formation and activation of the EEG. *Electroencephalogr. Clin. Neurophysiol*. **1**, 455–473. (doi:10.1016/0013-4694(49)90219-9)

141. Moller AR, Rollins PR. 2002 The non-classical auditory pathways are involved in hearing in children but not in adults. *Neurosci. Lett*. **319**, 41–44. (doi:10.1016/s0304-3940(01)02516-2)

142. Møller AR, Kern JK, Grannemann B. 2005 Are the non-classical auditory pathways involved in autism and PDD? *Neurol. Res*. **27**, 625–629. (doi:10.1179/016164105x25117)

143. van der Loo E, Congedo M, Vanneste S, De Heyning PV, De Ridder D. 2011 Insular lateralization in tinnitus distress. *Auton. Neurosci. Basic Clin*. **165**, 191–194. (doi:10.1016/j.autneu.2011.06.007)

144. Edlow BL *et al*. 2012 Neuroanatomic connectivity of the human ascending arousal system critical to consciousness and its disorders. *J. Neuropathol. Exp. Neurol*. **71**, 531–546. (doi:10.1097/nen.0b013e3182588293)

145. Jones BE. 2003 Arousal systems. *Front. Biosci. J. Virtual Libr* **8**, s438–451. (doi:10.2741/1074)

146. Sridharan D, Levitin D, Menon V. 2008 A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Networks* **105**, 12569–12574. (doi:10.1073/pnas.0800005105)

147. Uddin LQ. 2015 Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci*. **16**, 55–61. (doi:10.1038/nrn3857)

148. Nair J, Klaassen AL, Arato J, Vyssotski AL, Harvey M, Rainer G. 2018 Basal forebrain contributes to default mode network regulation. *Proc. Natl Acad. Sci. USA* **115**, 1352–1357. (doi:10.1073/pnas.1712431115)

149. Helmholtz H. 2009 *On the sensations of tone as a physiological basis for the theory of music*, 3rd edn. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511701801)

150. Oliveira G, Davidson A, Holczer R, Kaplan S, Paretzky A. 2016 A comparison of the use of glottal fry in the spontaneous speech of young and middle-aged american women. *J. Voice* **30**, 684–687. (doi:10.1016/j.jvoice.2015.08.015)

151. Tallon-Baudry C, Bertrand O. 1999 Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci* **3**, 151–162. (doi:10.1016/s1364-6613(99)01299-1)

152. Arnal LH, Wyart V, Giraud AL. 2011 Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci*. **14**, 797–801. (doi:10.1038/nn.2810)

153. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012 Canonical Microcircuits for Predictive Coding. *Neuron* **76**, 695–711. (doi:10.1016/j.neuron.2012.10.038)

13

royalsocietypublishing.org/journal/rstb   *Phil. Trans. R. Soc. B* **380**: 20240020