Review

Noises on—How the Brain Deals with Acoustic Noise

Livia de Hoz [1,2,*] and David McAlpine [1,3]

1   Neuroscience Research Center, Charité—Universitätsmedizin Berlin, 10117 Berlin, Germany
2   Bernstein Center for Computational Neuroscience, 10115 Berlin, Germany
3   Department of Linguistics, Macquarie University Hearing, Australian Hearing Hub,
     Sydney, NSW 2109, Australia
*   Correspondence: livia.dehoz@charite.de

Simple Summary: The listening brain must resolve the mix of sounds that reaches our ears into events, sources, and meanings. In this process, noise—sound that interferes with our ability to detect or understand sounds we need or wish to—is the primary challenge when listening. Importantly, noise to one person, or in one moment, might be an important sound to another, or in the next. Despite the many challenges posed by noise, however, human listeners generally outperform even the most sophisticated listening technologies in noisy listening environments. Because extracting noise from the sound stream is a fundamental process in listening, understanding how the brain deals with noise, in its many facets, is essential to understanding listening itself. Here, we explore what it is that the brain treats as noise and how it is processed. We conclude that the brain has multiple mechanisms for detecting and filtering out noise, and that incorporating cortico-subcortical 'listening loops' into our studies is essential to understanding this early segregation between noise and signal streams.

Abstract: What is noise? When does a sound form part of the acoustic background and when might it come to our attention as part of the foreground? Our brain seems to filter out irrelevant sounds in a seemingly effortless process, but how this is achieved remains opaque and, to date, unparalleled by any algorithm. In this review, we discuss how noise can be both background and foreground, depending on what a listener/brain is trying to achieve. We do so by addressing questions concerning the brain's potential bias to interpret certain sounds as part of the background, the extent to which the interpretation of sounds depends on the context in which they are heard, as well as their ethological relevance, task-dependence, and a listener's overall mental state. We explore these questions with specific regard to the implicit, or statistical, learning of sounds and the role of feedback loops between cortical and subcortical auditory structures.

Keywords: auditory; noise; background; foreground; statistical learning; feedback; loops; inferior colliculus; auditory cortex

1. The Challenge of Noise

> *'Nothing essential happens in the absence of noise.'* (Jacques Attali, French economist and philosopher)

Noise—which we define here as interfering sounds that mask what we are trying to hear or that divert our attention—is the primary challenge in listening. The listening brain must resolve the mix of sounds that reaches our ears into events, sources, and meanings that unfold over multiple cadences. Making sense of sound relies on separating what we want to hear—'signals'—from what we do not—'noise' [1]. Importantly, noise to one person, or in one moment, may be signal to another, or in the next. Processing noise (filtering it out or using its potential predictive power) is something the brain achieves seemingly effortlessly, but it is not obvious how. Normal-hearing listeners are adept at 'cocktail-party listening' despite the often very high levels of background noise, yet the cognitive effort involved

in this process is revealed in hearing-impaired listeners, who struggle to make sense of sound in even moderately challenging environments. Despite the many challenges noise presents, humans generally outperform even the most sophisticated listening technologies and are much better at extracting and parsing information and meaning from sounds that add context and 'texture' to listening.

And yet, we know little about what is classified as noise by the brain, and when so, and how and why the neural representation of noise is removed (the process of 'denoising') in the brain. Some sounds are always, or mostly, noise and rapidly relegated to 'background' (e.g., a waterfall, strong wind, a sudden downpour). Other sounds are, or become, 'background' (noise) when we wish to attend to 'foreground' sounds; a talker in a crowded room mentioning our name, perhaps. Loud or quiet, continuous or sporadic, embedded in the signal or generated by an entirely different source, the intensity, and texture of noise, all have a strong influence on how it is interpreted.

In this review, we address the question 'what is noise?' from perceptual and biological perspectives. We focus on those aspects of sound processing that enable us to determine what the auditory brain might at least consider noise. We discuss the effects of masking noise on perception and describe phenomena to illustrate how the brain has evolved to deal with noisy listening environments, including through binaural hearing. We then discuss how the brain adapts to background noise and how noise-invariant representations of sound might arise. Finally, we discuss how noise seems to automatically be incorporated into our perception of the world, and how it triggers sensory and motor actions to overcome its impact on our communication abilities. Two themes recur: that the brain contains mechanisms for detecting and dealing with noise implicitly—i.e., without engagement—and that subcortical structures are key to this process.

2. Noise as a Source of Interference—Energetic and Informational Masking

It is usually the case that noise of sufficient intensity harms listening performance by masking the signal we ought to hear (a predator approaching against strong wind in a forest). Indeed, sound intensity is a key determinant of how effective one sound is in masking another, with a winner-takes-all effect by which the energy of the signal determines its competitive strength, especially in the auditory periphery. Release from this 'energetic' masking can arise when the signal and the noise originate from different locations, a phenomenon known as spatial release from masking [2], and facilitated by binaural hearing (discussed below) or when the two signals hold different spectral compositions; for example, in terms of a talker's voice release [3]. When the masking sound is modulated, listening to the signal in the modulation minima—the trough of the noise signal or 'dip-listening'—can occur, and it is possible to follow a conversation if sufficient snippets of what a talker is saying are secured [4,5]. Nevertheless, even without dip-listening, the independence (from each other) of auditory channels provides a significant boost to following a conversation in external noise [6]. Perhaps not surprisingly, the specific nature of noise determines its masking capacity. Broadband noise, for example, despite its wide spectrum, is less detrimental to speech understanding compared to other forms of spectrally selective masking. Recent evidence suggests that the specific form of masking noise invokes very different brainstem, midbrain, and cortical circuits, even for the same level of speech understanding [7]. When listeners are engaged in the task of detecting non-words in a string of noisy words, vocoded speech—an intrinsically noisy representation—invokes brainstem circuits through the medial olivo-cochlear (MOC) reflex to sharpen cochlear filtering, supporting listening performance. For the same level of performance in the task, speech-shaped noise (noise with the same long-term spectrum of speech) elicited elevated levels of brainstem and midbrain auditory activity, whilst multi-talker babble—a particularly challenging form of masker—elicited the greatest level of cortical activity, including processing indicative of greater cognitive engagement. These differences were less evident, or altogether absent, during passive listening, suggesting that the listening

brain is equipped with multiple mechanisms—from ear to cortex and back—for dealing with different forms of noise and different mental states.

Informational Masking

Beyond energetic masking, the concept of 'informational' masking has been used to describe situations in which non-overlapping sound energy nevertheless impairs (masks) speech understanding (see review in [8]). Intuitively, this makes sense; hearing a conversation against a background of other talkers, for example, likely requires active engagement in listening and the use of cognitive resources. Even when the acoustic energy of two talkers is largely non-overlapping, information conveyed by the unattended talker can impact that conveyed by the attended; it becomes noise for the purposes of understanding. Nevertheless, defining noise as energetic (overlapping) <u>or</u> informational (non-overlapping) is perhaps a simplification. Neural activity generated by any sound, even in the auditory periphery brainstem, is not instantaneous. For example, evidence that a time-reversed speech signal generates less informational masking than forward-running speech [9] assumes that masking relates to semantic meaning—reversed speech (noise) has the same energy as the signal (forward speech) but, being unintelligible, is less effective as an informational masker. However, natural speech also contains a very different temporal waveform to reversed speech—a series of rapid acoustic attacks—favored by biophysical mechanisms responsible for neural firing—and slower decays, whilst the temporal waveform of reverse speech shows the opposite pattern—slow rises in energy and rapid decays—likely eliciting fewer and less-precise neural responses [10]. Beyond acoustics, but well before semantics, forward and reverse speech elicit different amounts of 'neural' energy. *In extremis*, non-overlapping acoustic transients (clicks) can be rendered temporally inseparable by the 'ringing' response of the basilar membrane (see the discussion of the precedence effect, below), rendering their overlapping status moot. Whilst energetic and informational masking are useful concepts in describing acoustics and perception, their neural counterparts are currently less well-conceptualized.

3. Binaural Hearing—The Brain's Denoising Algorithm

Though many factors likely contribute to successful listening in noise, including a healthy inner ear, it is indisputably the case that binaural, or two-eared, hearing is key to successful 'cocktail-party listening', the ability to follow a conversation against a background of noise, competing voices, or reverberant sound energy [11]. Possessing two functional ears provides us with an enormous listening benefit beyond the 3-decibel benefit provided by two independent sound receivers or even the ability to locate sources of sound based on sensitivity to binaural cues—interaural time and level differences (ITDs and ILDs, respectively).

The most well-known and well-studied of the binaural benefits is binaural unmasking (see [12] for a comprehensive review), an improvement in the detection of sounds or the intelligibility of speech in background noise based on the relative interaural configurations of the signal and masker (noise). Independently reported in 1948 [13,14], the ability to hear out sounds in background noise by inverting the fine-structure phase of the low-frequency signal (<1500 Hz) or masker waveform at one ear relative to the other has spawned an entire research field and supports a wide range of audio technologies, listening devices, and therapeutic interventions (Figure 1A). The remarkable denoising capacity of binaural hearing is demonstrated most compellingly in that it overrides the addition of sound energy *per se* to accrue a listening benefit. Specifically, if a tone and a fixed level of masking noise are presented monaurally to the same ear and the level of the tone reduced to the point it becomes undetectable, the tone becomes detectable again simply by adding identical noise to the other ear. Its level must be reduced once more to reach the 'masked threshold'. Remarkably, adding an identical noise to the other ear makes it easier once more to hear the signal. Then, by adding an identical tone to the other ear—such that both signal and noise are identical across the ears—the tone becomes more difficult to detect, despite the overall

increase in signal energy across the ears—necessitating an increase in level to perceive it. Finally, by inverting the phase of the tone or the noise in one ear relative to the other, the tone becomes audible again, more so when the tone is inverted—the classic binaural unmasking paradigm (Figure 1A). This counterintuitive effect relies on brain mechanisms that compare the relative phases of the signal and the masker at each ear, and the relative difference in signal and masker phase across the ears determines the magnitude of the unmasking benefit. Clearly evolved to support real-world listening in noisy environments, binaural unmasking combines with the effect of head shadowing and perhaps even the absolute sensitivity to monaural sound levels to allow spatial release from masking—the ability to hear sounds based on the relative location of specific sources and interfering noise.
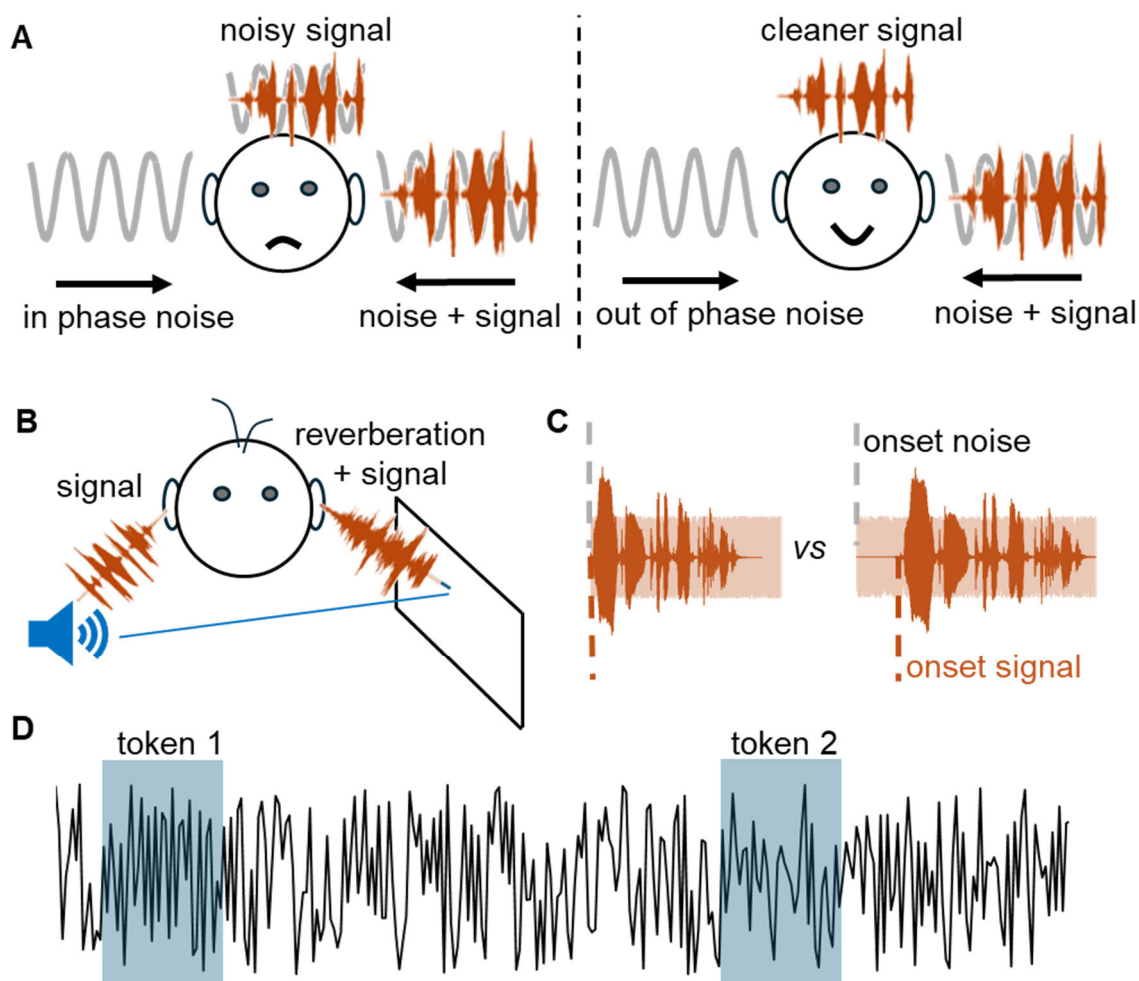
Figure 1. Schematics of noise effects in different contexts. (**A**) Binaural unmasking as a result of out-of-phase binaural noise. On the left, the noise coming into the left and right ears is in phase such that the signal (copper) coming into the left ear is only partially denoised by the brain, as shown by drawing above the head. On the right, the noise coming into the left and right ears is out of phase and the signal (copper) is perceived more cleanly by the brain, as shown by the drawing above the head. Inspired by [15]. (**B**) Hearing in reverberating environments. The sound (copper) arrives directly from the source (blue speaker) into the right ear and indirectly, as a reflection from the wall and contaminated by reverberation, into the left ear. (**C**) The advantage of pre-exposure to background sound. On the left, the vocalization (copper wave) and the background noise (pink area behind) start at the same time. On the right, the vocalization starts later, a condition that facilitates its understanding. (**D**) Textures vs. exemplars. The specific statistics of two short tokens of sound (1 and 2) can differ from one another (token 1 contains higher frequencies than token 2) and from the summary statistics of a long texture sound.

Neural Mechanisms for Binaural Unmasking

A neural basis for binaural unmasking was established in a series of in vivo experiments confirming the existence of neurons in the midbrain—likely innervated by inputs from primary binaural brainstem neurons—that demonstrate features consistent with the perceptual effect and existing models [16–18]. Specifically, neurons in the inferior colliculus sensitive to ITDs show lower thresholds to out-of-phase binaural tones than to in-phase binaural tones in binaural background noise [18]. Brain mechanisms contributing to this ability are likely highly conserved across species, with an exception being fish [19]. To this point, a series of experiments across species, brain centers, and recording modalities by Dietz and colleagues demonstrated that the auditory brain's capacity to deal with more realistic types of sound in reverberant environments relies on fast temporal capacity of binaural processing, facilitating the extraction of reliable spatial cues arriving direct from the source and suppressing before the later arrival of more- energetic, reverberant sound energy in which spatial cues are scrambled [20,21]. Importantly, this binaural form of denoising relies on monaural inputs that are already strongly adapted at the level of the auditory nerve; modeling data [22] suggest that human speech sounds in reverberation are consistently mis-localized in the absence of monaurally adapting inputs to binaural neurons. Brain mechanisms supporting binaural hearing seem highly evolved to cope with noisy and reverberant environments, with access to temporal information conveyed in low-frequency sounds particularly important in achieving this.

4. Listening Spaces as a Source of Noise—Dealing with Reverberation

One common source of noise we often must deal with is reverberant sound energy, or echoes (Figure 1B). Most natural, including open, spaces generate acoustic reverberations [23], and these can provide useful information as to the size, nature, purpose and importance of the spaces we inhabit [24]. Often, the direct-to-reverberant ratio (DRR) can be negative, i.e., sound energy direct from the source can be less intense than later-arriving reflections and interference patterns, yet normal-hearing listeners seem able to parse these complex environments with relative ease. Even when absolute listening performance is impaired by reverberation, our brains have evolved to deal with it to some degree. For example, listeners seem able to adapt to the reverberant characteristics of rooms [25] with just a few milliseconds of prior listening exposure to the sounds in a room supporting better speech understanding (Figure 1C) [26]. This benefit does not accrue in anechoic environments, suggesting that parsing speech in noisy, reverberant environments taps into brain mechanisms that implicitly learn the acoustic features of background sounds to improve explicit processing of foreground ones. Rather than the now common act of applying denoising algorithms to remove (presumed) external noise though listening devices ('hearables', earbuds, and the like), listeners might benefit from being pre-adapted to noisy and reverberant listening environments in which they need or wish to communicate [1]. Research fields such as eco-acoustics have developed in acknowledgement of the fact that listening is optimized for real-world spaces. This becomes more important as the world becomes noisier, with initiatives taking place to try to preserve natural soundscapes, for example, the 'One square inch of silence' [27]. Individual listeners are often ignored in our listening spaces, with other sources of 'noise'—cognition, hearing status, age, language, and culture—rarely featuring in their design, construction, or assessment.

The Precedence Effect—Direct Is Correct

Given our need to navigate the spaces we inhabit, reverberation also generates ambiguities in terms of the perceived location of sound sources. In particular, despite the presence of multiple reflected copies of the sound arriving from different locations milliseconds later, listeners usually report a single sound image originating from its source. This dominance of directional cues conveyed by earlier-arriving information has been explored through a set of phenomena collectively referred to as the precedence effect (sometimes referred to as the 'Haas' effect [28]). In one common paradigm used to explore precedence,

two identical sounds are presented from two different speakers, whether simultaneously or with a delay below the so-called echo threshold, so they will be heard as a single sound ('fusion'). As the delay is increased further, the leading sound dominates the perceived location until, for delays that are longer still, two, spatially distinct, sources are perceived. Together, localization dominance and fusion permit more accurate processing of sources in acoustically reverberant environments [29]. In vivo physiological studies of the precedence effect report reduced neural responses to lagging sounds not explicable purely by neural adaptation [30,31]. Neural correlates have been suggested at the level of the inferior colliculus [32] (see also ref. [33]), in which neural responses to lag stimuli are suppressed up to 40 ms after the presentation of a lead stimulus [34]. The magnitude and time course of neural responses to lagging clicks, which depend on the ITDs imposed on leading and lagging clicks, have been correlated with behavioral measures of precedence [30,31,33], suggesting an ITD-dependent inhibition that elicits a temporary break in interaural sensitivity [35]. This neural-suppression hypothesis assumes neural responses to each binaural click are processed separately and in sequence. Despite widespread acceptance of this view, however, an alternate hypothesis suggests the psychophysical effect is explicable without recourse to neural inhibition or, indeed, any central auditory processing. Instead, the finite response time of the basilar membrane to transient stimuli, especially at the apical, low-frequency end of the cochlea, renders leading and lagging sounds a single event on the basilar membrane [36,37]. Depending on the interval between them, the amplitude and timing cues conveyed by leading and lagging clicks will be modified such that, following binaural integration, the internal representation of directional cues differs from that presumed from the external stimuli. For species that exploit directional cues present in the low-frequency components of sound, localization judgments would be determined by these altered cues.

Regardless of the precise mechanisms contributing to precedence, the 'echo-threshold' can vary substantially depending not only on the type of sound (up to about 5 ms for clicks and up to 40 ms for speech or other complex sounds) but also on the environment. For example, in anechoic and reverberant environments, sounds arriving from opposite sides will be fused for delays of up to about 5 and over 30 ms, respectively [38,39]. Broadband noise superimposed over this time interval will have little effect on these thresholds, again supporting the notion that broadband noise, unless overly intense, is easily filtered out by the brain.

Overall, whether through mechanical or computational means, the logic of physics prevails—sound arriving direct from a source is likely to arrive earlier than its echo(es) and the brain 'accepts' this implicitly.

5. Filling in the Gaps—Noise as a Masker Even When It Is Not

Although our brains evolved in natural soundscapes and environments, synthetic and illusory sounds can also be informative as to the brain's function and its implicit sensitivities. This is particularly true of illusions that result in the preservation (or creation) of stable percepts. One such illusion of relevance to how the brain deals with noise is the 'continuity illusion'. Tones presented as interrupted pulses are perceived as such until the silent gaps in between are filled with noise. This simple manipulation transforms the sound percept from one of disconnected tones to one of a single, continuous tone pulse. This illusion is stable across species, including humans [40–44], non-human primates [45–47] and cats [48]. It is also resilient to the type of sound, from modulated tones [41,42,49], music [50], and vowels [51], to more complex speech stimuli [52]. Interestingly, the illusion only occurs when the noise fills the gap, not when the noise is presented continuously over both the gap and sounds [40,42,44,45]. While the mechanisms are not understood, neurons in the primary auditory cortex of cats and monkeys respond to the illusion as if it was a continuous sound, but not to the individual tones, nor to the noise alone [46,48]. Once more, broadband noise is automatically blended into the background.

6. Responding Reflexively to Noise—The Lombard Effect

The brain's adaptations to communication in noise extend beyond those that enhance the signal-to-noise ratio of what we hear to increase the intensity of what we emit. When background sound intensities increase, animals respond by automatically and involuntarily vocalizing more intensely. In humans, this 'Lombard reflex' (named for its discoverer) extends to speech properties such as the rate and duration of syllables [53]. Whilst the Lombard reflex evolved to be dynamic in environments with fluctuating intensities, the overall noisier urban spaces of the 21st century affect how we and other animals vocalize. This became evident during the COVID-19 pandemic, for example, when birds in the San Francisco Bay area reduced the intensity of their song, presumably as a response to the measurable reduction in the intensity (and occurrence) of environmental sounds during the period of enforced lockdown [54].

The Lombard reflex has been observed in numerous vertebrate species from fish to birds and mammals, with few exceptions [55], and thus has a possible evolutionary origin dating as far back as 450 million years. While the mechanisms remain unknown, the fact that it is evident in decerebrate cats and fish supports the importance of subcortical structures in its emergence. Its rapid latency relative to intensity increases also supports a role for subcortical circuits. While latencies of ~150 ms have been measured in humans and other animals [56–58], in bats the increase in the call amplitude can occur as fast as 30 ms after sound onset [59], too soon for higher-order brain structures to be involved or feedback on the intensity of self-vocalizations to be assessed. The effect tracks not only upward but also downward changes in the background intensity, so long as these occur over time windows of >50 ms in bats [59], suggesting temporal summation on time scales of tens of milliseconds. Two lines of evidence suggest that cortical influences support the emergence and magnitude of the effect. While the Lombard reflex is normally automatic, subjects can learn to control it using, for example, visual feedback on the loudness of their voices [60]. Another clue to the circuit mechanisms can be derived from the frequency-specificity of the reflex. While most studies use broadband noise to induce the reflex, using narrower band noise has revealed, in birds, bats, humans and monkeys, that increases in amplitude only occur when the noise covers the spectral band of the vocalization [61–64]. Thus, the circuit shows some precision in its tonotopic specificity. Whether this is learned or hard-wired is not known.

7. Stochastic Noise and the Brain's Internal State

It is worth noting that some noise is internal, and not all of it is bad. Within the inner ear, 'stochastic' noise—random fluctuations in input—such as Brownian motion in the mechano-transducing sensory hair cells, far from harming performance, improves their sensitivity, whilst the high (often $>100 \text{ s}^{-1}$) spontaneous firing rates of auditory nerve fibers render them independent of each other, enhancing the flow of information to the listening brain [65]. Stochastic noise might even expand the information bottleneck in cochlear implants [66] and could inform new signal-processing strategies that convey a richer listening experience in listening devices [67].

Despite the apparent utility, however, <u>elevated levels of</u> internal noise are generally considered a form of pathology, with maybe the most relevant of these being tinnitus—the perception of sound in the absence of any external sound source. The phenomenon of tinnitus illustrates the delicate balance in the adult brain between external input and internal activity and how a disconnection between the two might have the effect of generating unwanted internal noise in the system beyond a mere reduction in sensitivity. Affecting maybe 10% of the population, and afflicting half of these, tinnitus (or 'ringing in the ears') is long associated with damage to the sensory hair cells responsible for our sensitive hearing [68]. So-called *objective* tinnitus can be understood as the result of compensatory, homeostatic upregulation of neural activity (or down-regulation of neural inhibition) in the central nervous system, perhaps to maintain some required long-term level of neural responsiveness. This maladapted neural gain generates elevated neural 'noise'—neural

firing not related to external stimulus—with the concomitant perceptual experience of sound. This explanation once fell short when trying to account for *subjective* tinnitus, tinnitus that arises without any obvious sign of inner ear or brain pathology. Recently, however, the concepts of cochlear synaptopathy [69]—pathologies to synaptic transmission in a specific population of high-threshold inner-ear nerve fibers that code for higher sound intensities—suggests an occult form of hearing deficit (coining the term 'hidden hearing loss' [70]) that manifests as tinnitus as well as problems listening in background noise ('I can hear you but I can't understand you') in otherwise normal-hearing listeners. This perspective is supported by in vivo research demonstrating that neural coding of speech in background noise [71] and the ability to adapt to different noisy environments [72] are both impaired by synaptopathic insults to the inner ear that spare hearing thresholds. A recent theoretical report suggests that feedforward adaptive changes in the level of internal stochastic noise arising from hearing loss—including 'hidden'—can explain phantom percepts such as tinnitus in the context of the Bayesian brain framework [73]. In this perspective, a mismatch between the brain's *a priori* expectations and the input results in neural noise being interpreted as signal. An important facet of this construct is the brain's expectation of a <u>familiarized</u> level of internal, stochastic noise along the auditory pathway, and a resulting pathology of a phantom percept when the level of internal noise changes.

8. From Foreground to Background—And Back

Whether foreground or background, signal or noise, all the sounds processed by the inner ear likely impact the auditory brain, a key distinction being the extent to which they are consciously experienced (foreground more than background). Given the great many sounds that might impinge on us at any one time, there is a clear advantage to processing many of these subconsciously; background sounds representing a form of 'aural wallpaper' to be accessed more fully when it becomes important to do so. A relatively recent perspective on the issue of how this is achieved takes a more ethological approach than traditional psychoacoustics in terms of considering how the listening brain deals with the complexities of cluttered acoustic scenes—the classic Bregmanian 'auditory scene analysis' problem [40,44]. These studies are often predicated on the fact that (1) whilst foreground sounds must be attended to, background sounds—'noise'—must also be processed in order to provide access to the entire auditory scene, including to sounds that might soon become foreground or that represent a lurking danger, perhaps, and (2) that the brain has evolved specifically to deal with this problem, particularly over time. Lines of investigation explore how the auditory system detects order (signal) from disorder (noise)—evidence, perhaps, of a source emerging from the background. Employing sequences of tones that transition between random and regular patterns, these studies demonstrate a clear difference in processing time and the associated neural signatures depending upon whether a transition is from random to regular or vice versa [74]. Specifically, determining that a sequence contains some form of acoustic structure takes longer to process than departures from regularity once knowledge of regularity is established [75], and this process is more affected by interfering acoustic events than visual ones [76]. Another way of exploring this issue concerns the extent to which 'textures'—naturalistic background sounds such as fire, water, wind, rain etc.—are represented in the brain not as objects *per se* but in terms of their summary statistics, at least for sounds of sufficient duration [77]. Again, there is a clear temporal dimension to this type of processing; short 'tokens' of sound textures are perceived as objects, distinguishable from other tokens drawn from the same statistical pool. With increasing duration, however, individual tokens of textures are less distinctive from each other, and eventually indistinguishable, whilst sensitivity to their longer-term 'summary' statistical structure increases. This suggests a continuum over which foreground and background sounds are not distinct, but rather are processed as if they exist along the same temporal and statistical dimensions. Sound objects emerge from noise, or fade into it, depending on their positions along these dimensions.

However, although the ability to process background sounds absent obvious conscious engagement is likely of evolutionary advantage, a potential confounder in these studies is the extent to which listeners can suppress seemingly innate abilities to impose some form of structure on them or meaning to them other than the stimulus dimension presumed in the experimental paradigm. Specifically, though expressed in terms of the automatic, potentially subconscious, neural processing of sounds, these studies, by requiring subjects to make explicit distinctions between noise bursts [78], sound textures [77], or streams of tones [74], 'force' supposedly background sounds to the foreground. Furthermore, for noise bursts and sound textures at least, casting listening tasks in terms of perceptual sensitivity to summary statistics can require the removal of most (up to 95%) of exemplars on the basis that they are insufficiently 'statistical' in nature (i.e., they contain potential sound features to which listeners might be sensitive). If abstract or natural sounds cannot simply be condensed to their summary statistics, perhaps the main task of the listening brain is constantly to suppress the perception of spurious events and sound objects rather than straining to extract them in the first place.

Dichotic Pitches—Creating Sounds Objects from Noise

Given the propensity of human listeners to generate the percept of sound objects from seemingly unstructured and random sounds, including noise (auditory pareidolia, often considered in terms of hallucinations; [79]), it is perhaps unsurprising that trying to account for performance in listening tasks directed toward supposedly background sounds (i.e., turning the background to foreground) remains fraught. To this end, it is worth asking what minimum sound features are required to generate a percept of foreground sounds (objects) against a background of noise. Though seldom expressed this way, the concept of dichotic pitches is telling. Dichotic pitches are clear percepts of sound objects generated when otherwise interaurally incoherent noise (independent samples of noise presented to either ear simultaneously) is briefly transitioned to coherent by applying deterministic phases and magnitudes to the signal (either the whole signal or at specific frequencies, [80]). Like sound textures, very short snippets of incoherent noise are perceived as punctate sound objects holding a distinct intracranial location determined by their combination of interaural time and level disparities. This punctate, and localizable, percept is lost, however, as the stimulus duration is increased, to be replaced by a relatively broad intracranial percept of inchoate noise [81,82]. Applied in different frequency bands over time, transitions from incoherent to coherent noise are sufficiently robust to generate melodies, distinguishable against an otherwise indistinct, noisy background [83]. This is despite the phase and magnitude transitions being imperceptible at either ear alone. Based on our knowledge of binaural hearing, the neural signature for dichotic pitches likely arises first post-synaptically in the superior olive, three synaptic stages beyond the inner ear, demonstrating the listening brain's capacity to extract features from background noise even in the absence of any externally imposed modulation in sound energy.

9. Brain Circuits and Listening Loops

In seeking to understand how noise is represented in the brain, one common approach has been to study the extent to which the neural representation is invariant to noise. Such a representation, which would favor a neural representation of the signal, would be evidence for a brain mechanism of background suppression. Typically, a foreground signal, speech for example, is presented in a background of noise of different intensities, generating varying signal-to-noise ratios (SNRs). The study of noise-invariant sparse representations has focused on cortical structures and indeed neurons across sensory modalities [84] and, within the auditory modality, across species [85–89], have been found to show noise-invariance. Cortical noise-invariant representations are not incompatible with representations of noise itself in the same structures. For example, for noises of sufficient intensity, a substantial fraction of neurons in the cortex of the cat responded to a bird chirp embedded in noise in the same way they would to the noise alone, despite

the chirp being considerably louder than the noise [90]. These data reiterate the question of 'when is noise noise and when is it signal?', and raises the issue of acoustic saliency. Does background noise, when sufficiently intense, impose itself to become part of the foreground? These studies often use speech or vocalizations as the signal embedded in naturalistic noises. It is possible that the use of a high-level signal anchors the invariance to higher cortical regions where speech is encoded. In any case, it is not unreasonable to think that for a noise-invariant representation to arise, the detection and subtraction of noise-evoked or -associated neural responses might have taken place upstream in the pathway. Evidence of just this has been demonstrated at the level of the cochlear nucleus in mammals [91] and the avian midbrain [92]. That noise might gradually be extracted from the signal along the auditory hierarchy is supported by electrophysiological studies [88] reporting increased noise invariance between the inferior colliculus and the auditory cortex, and from theoretical considerations [93]. Nevertheless, some across-hierarchy comparative studies report that the brain centers showing the most noise-invariant representation are the auditory thalamus and inferior colliculus [94,95] rather than cortex. In an anesthetized guinea pig, the presence of noise reduces neural discrimination of vocalization tokens across all the structures from the cochlear nucleus to the primary auditory cortex, probably because of the attenuation of slow amplitude-modulated cues that provide the temporal envelope for the test vocalizations. Interestingly, the inferior colliculus and, to a lesser extent, the thalamus provided the best discrimination level in masking noise [95]. The authors extended these results to vocalization presented in stationary or chorus (colony of vocalizing guinea pigs) noise and again the inferior colliculus and thalamus represented vocalizations with the highest fidelity [94], while neurons in the cochlear nucleus displayed the most faithful representation of the noise itself. Increasingly, the cortex is explored in terms of it being a source of feedback instruction to subcortical centers and, in this regard, inactivation of cortex bilaterally (through cooling) in an anaesthetized guinea pig abolishes the extent to which neurons in the inferior colliculus 'learn' over time the statistical distribution of the intensities of noise to speed up their adaptive capacity [96].

Nevertheless, it is not a question of whether auditory neurons respond to tokens of noise but of at what level of processing is the presence of background noise detected and the response to it extracted, especially when it forms a continuous background (energetically) masking potentially relevant signals. Context plays a crucial role in the interpretation of noise. For example, whether we interpret an intense background noise as an innocuous ventilation system, or a potentially faulty machine will depend on where we are and our memory of it. Others have implicated the auditory cortex, both primary and higher order [97,98], in contextual learning, albeit inferring contextual learning from its effect on tasks such as speech recognition. Subcortical structures are sensitive to the statistics of contextual sound though. Predictable, but not unpredictable, contextual signals generate more acute neural representations of sound by neurons in the inferior colliculus [99,100].

Cortico-Subcortical Loops

A complex process such as contextual processing is likely to occur simultaneously at several interacting levels of the auditory system in a time- and task-dependent manner, possibly over progressively longer time windows [101,102]. The concept that interactions between cortical and subcortical structures might play a key role in active sensing, the flexible processing of information depending on the context and task at hand, is becoming progressively more present in the neuroscience community [100,103,104]. Corticofugal projections, which allow for cortical feedback, have been implicated in flexibility across modalities [105] and, in the auditory system, reach all the upstream levels, including the cochlear nucleus and the inferior colliculus. Indeed, in the context of noise, the cortex helps adjust the response gain to the stimulus intensity through down-regulation of the cochlear gain in loud environments, i.e., the cochlea in the case of the auditory system [106]. The inferior colliculus receives feedback projections from deep layers of the primary auditory cortex [107,108] and there is in fact substantial evidence of a role of these cortico-collicular

projections in learning [109–111]. Indeed, inhibition of the cortico-collicular projection results in delays in learning [112] and response adaptation [96], and it prevents fast escape responses to loud sounds [113]. Understanding the role of these projections and the circuit involved will help us understand the interactions between background and foreground.

One example of such an interaction is the way in which the brain deals with the statistical structures of background noise. The intensity of background noise can vary rapidly over a wide range and auditory neurons cope by adapting their sensitivity curves to match the intensity statistics of the current context [96,114] and the extent to which neurons adapt to the statistical distribution of sound intensities increases from the auditory nerve [115] through the midbrain [114,116] to the cortex [117]. The response functions of neurons in the inferior colliculus shift to match the current distribution of sound intensities [116], and they do so with a time course suggesting that they are adapted to account for ecologically relevant sounds [118]. While we do not yet understand the time windows of integration used to detect these distributions, or even the mechanisms of this integration, we know that the dynamics of this adaptation are encoded in the inferior colliculus over relatively long time courses, inform adaptation in the cortex [118], and depend on cortico-collicular interactions [96], demonstrating the importance of listening loops in setting the cadences over which noise is assessed and dealt with by the auditory brain.

10. Conclusions

We have provided an overview of different perceptual and biological phenomena that demonstrate that the brain inherently expects and incorporates noise into its listening experience. It sustains mechanisms to deal with noise as a masker of signals, both to reduce its impact and to infer what lies beneath (e.g., continuity illusion). This suggests that dealing with noise is a core feature of the listening brain. Understanding how the listening brain separates signal from noise requires a combination of biological, psychophysical, and theoretical approaches and a bridge—in the form of transformative experimental design—between in vivo models and human listeners. Spanning scales of magnitude, complexity, and species, we can generate hypotheses tailored for exploration in human and animal listeners to determine how the statistics of the environmental sound, on the one hand, and the arousal state, attention, or goal-oriented behaviors, on the other, impact the ability to parse complex acoustic scenes. In awake, freely moving animals, we can establish fundamental links between brain activity and listening performance, revealing critical processes and circuits responsible for the brain's ability to make sense of cluttered sound environments, but we are still far from understanding the listening brain's remarkable and implicit 'denoising' capacity, including its ability to shape the sensitivity of the ear itself to enhance listening performance. This is necessary to establish clear links back to data and theories developed in animal models, instantiating an iterative approach to understanding the brain's capacity for effective listening in noise.

Author Contributions: Conceptualization, L.d.H. and D.M.; investigation, L.d.H. and D.M.; resources, L.d.H. and D.M.; writing—original draft preparation, L.d.H. and D.M.; writing—review and editing, L.d.H. and D.M.; project administration, L.d.H. and D.M.; funding acquisition, L.d.H. and D.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Haven Wiley, R. *Noise Matters: The Evolution of Communication*; Harvard University Press: Cambridge, MA, USA, 2015.
2. Freyman, R.L.; Helfer, K.S.; McCall, D.D.; Clifton, R.K. The Role of Perceived Spatial Separation in the Unmasking of Speech. *J. Acoust. Soc. Am.* **1999**, *106*, 3578–3588. [CrossRef] [PubMed]
3. Freyman, R.L.; Helfer, K.S.; Balakrishnan, U. Spatial and Spectral Factors in Release from Informational Masking in Speech Recognition. *Acta Acust. United Acust.* **2005**, *91*, 537–545.
4. Moore, B.C.J. The Importance of Temporal Fine Structure for the Intelligibility of Speech in Complex Backgrounds. *Proc. ISAAR* **2011**, *3*, 21–32.
5. Vélez, A.; Bee, M.A. Dip Listening and the Cocktail Party Problem in Grey Treefrogs: Signal Recognition in Temporally Fluctuating Noise. *Anim. Behav.* **2011**, *82*, 1319–1327. [CrossRef]
6. Apoux, F.; Healy, E.W. On the Number of Auditory Filter Outputs Needed to Understand Speech: Further Evidence for Auditory Channel Independence. *Hear. Res.* **2009**, *255*, 99–108. [CrossRef] [PubMed]
7. Hernández-Pérez, H.; Mikiel-Hunter, J.; McAlpine, D.; Dhar, S.; Boothalingam, S.; Monaghan, J.J.M.; McMahon, C.M. Understanding Degraded Speech Leads to Perceptual Gating of a Brainstem Reflex in Human Listeners. *PLoS Biol.* **2021**, *19*, e3001439. [CrossRef]
8. Farrell, D. Auditory Informational Masking—Gerald Kidd Jr. And Christopher Conroy. Acoustics Today. Available online: https://acousticstoday.org/auditory-informational-masking-gerald-kidd-jr-and-christopher-conroy/ (accessed on 21 June 2024).
9. Kidd, G., Jr.; Best, V.; Mason, C.R. Listening to Every Other Word: Examining the Strength of Linkage Variables in Forming Streams of Speech. *J. Acoust. Soc. Am.* **2008**, *124*, 3793–3802. [CrossRef]
10. Pressnitzer, D.; Winter, I.M.; Patterson, R.D. The Responses of Single Units in the Ventral Cochlear Nucleus of the Guinea Pig to Damped and Ramped Sinusoids. *Hear. Res.* **2000**, *149*, 155–166. [CrossRef] [PubMed]
11. Cherry, E.C. Some Experiments on the Recognition of Speech, with One and with Two Ears. *J. Acoust. Soc. Am.* **1953**, *25*, 975–979. [CrossRef]
12. Colburn, H.S.; Durlach, N.I. Models of Binaural Interaction. *Handb. Percept.* **1978**, *4*, 467–518.
13. Licklider, J. The Influence of Interaural Phase Relations upon the Masking of Speech by White Noise. *J. Acoust. Soc. Am.* **1948**, *20*, 150–159. [CrossRef]
14. Hirsh, I. The Influence of Interaural Phase on Interaural Summation and Inhibition. *J. Acoust. Soc. Am.* **1948**, *20*, 536–544. [CrossRef]
15. Moore, B.C.J. *An Introduction to the Psychology of Hearing*, 5th ed.; Academic Press: San Diego, CA, USA, 2003.
16. Palmer, A.R.; Jiang, D.; McAlpine, D. Neural Responses in the Inferior Colliculus to Binaural Masking Level Differences Created by Inverting the Noise in One Ear. *J. Neurophysiol.* **2000**, *84*, 844–852. [CrossRef] [PubMed]
17. Palmer, A.R.; Jiang, D.; McAlpine, D. Desynchronizing Responses to Correlated Noise: A Mechanism for Binaural Masking Level Differences at the Inferior Colliculus. *J. Neurophysiol.* **1999**, *81*, 722–734. [CrossRef] [PubMed]
18. Jiang, D.; McAlpine, D.; Palmer, A.R. Responses of Neurons in the Inferior Colliculus to Binaural Masking Level Difference Stimuli Measured by Rate-versus-Level Functions. *J. Neurophysiol.* **1997**, *77*, 3085–3106. [CrossRef] [PubMed]
19. Veith, J.; Chaigne, T.; Svanidze, A.; Dressler, L.E.; Hoffmann, M.; Gerhardt, B.; Judkewitz, B. The Mechanism for Directional Hearing in Fish. *Nature* **2024**, *631*, 118–124. [CrossRef]
20. Dietz, M.; Marquardt, T.; Stange, A.; Pecka, M.; Grothe, B.; McAlpine, D. Emphasis of Spatial Cues in the Temporal Fine Structure during the Rising Segments of Amplitude-Modulated Sounds II: Single-Neuron Recordings. *J. Neurophysiol.* **2014**, *111*, 1973–1985. [CrossRef] [PubMed]
21. Dietz, M.; Marquardt, T.; Salminen, N.H.; McAlpine, D. Emphasis of Spatial Cues in the Temporal Fine Structure during the Rising Segments of Amplitude-Modulated Sounds. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15151–15156. [CrossRef] [PubMed]
22. Brughera, A.; Mikiel-Hunter, J.; Dietz, M.; McAlpine, D. Auditory Brainstem Models: Adapting Cochlear Nuclei Improve Spatial Encoding by the Medial Superior Olive in Reverberation. *J. Assoc. Res. Otolaryngol.* **2021**, *22*, 289–318. [CrossRef]
23. Traer, J.; McDermott, J.H. Statistics of Natural Reverberation Enable Perceptual Separation of Sound and Space. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E7856–E7865. [CrossRef]
24. Blesser, B.; Salter, L.-R. *Spaces Speak, Are You Listening: Experiencing Aural Architecture*; MIT Press: Cambridge, MA, USA, 2009.
25. Srinivasan, N.K.; Zahorik, P. Prior Listening Exposure to a Reverberant Room Improves Open-Set Intelligibility of High-Variability Sentences. *J. Acoust. Soc. Am.* **2013**, *133*, EL33–EL39. [CrossRef] [PubMed]
26. Brandewie, E.J.; Zahorik, P. Speech Intelligibility in Rooms: Disrupting the Effect of Prior Listening Exposure. *J. Acoust. Soc. Am.* **2018**, *143*, 3068. [CrossRef] [PubMed]
27. Hempton, G.; Grossmann, J. *One Square Inch of Silence: One Man's Quest to Preserve Quiet*; Free Press: New York, NY, USA, 2010.
28. Haas, H. The Influence of a Single Echo on the Audibility of Speech. *J. Audio Eng. Soc.* **1972**, *20*, 146–159.
29. Wallach, H.; Newman, E.B.; Rosenzweig, M.R. The Precedence Effect in Sound Localization. *Am. J. Psychol.* **1949**, *62*, 315–336. [CrossRef] [PubMed]
30. Yin, T.C. Physiological Correlates of the Precedence Effect and Summing Localization in the Inferior Colliculus of the Cat. *J. Neurosci.* **1994**, *14*, 5170–5186. [CrossRef] [PubMed]
31. Fitzpatrick, D.C.; Kuwada, S.; Batra, R.; Trahiotis, C. Neural Responses to Simple Simulated Echoes in the Auditory Brain Stem of the Unanesthetized Rabbit. *J. Neurophysiol.* **1995**, *74*, 2469–2486. [CrossRef]

32. Litovsky, R.Y.; Delgutte, B. Neural Correlates of the Precedence Effect in the Inferior Colliculus: Effect of Localization Cues. *J. Neurophysiol.* **2002**, *87*, 976–994. [CrossRef]

33. Litovsky, R.Y.; Yin, T.C. Physiological Studies of the Precedence Effect in the Inferior Colliculus of the Cat. I. Correlates of Psychophysics. *J. Neurophysiol.* **1998**, *80*, 1285–1301. [CrossRef]

34. Tollin, D.J.; Populin, L.C.; Yin, T.C.T. Neural Correlates of the Precedence Effect in the Inferior Colliculus of Behaving Cats. *J. Neurophysiol.* **2004**, *92*, 3286–3297. [CrossRef] [PubMed]

35. Zurek, P.M. The Precedence Effect and Its Possible Role in the Avoidance of Interaural Ambiguities. *J. Acoust. Soc. Am.* **1980**, *67*, 953–964. [CrossRef]

36. Tollin, D.J.; Henning, G.B. Some Aspects of the Lateralization of Echoed Sound in Man. I. The Classical Interaural-Delay Based Precedence Effect. *J. Acoust. Soc. Am.* **1998**, *104*, 3030–3038. [CrossRef] [PubMed]

37. Hartung, K.; Trahiotis, C. Peripheral Auditory Processing and Investigations of the "Precedence Effect" Which Utilize Successive Transient Stimuli. *J. Acoust. Soc. Am.* **2001**, *110 Pt 1*, 1505–1513. [CrossRef] [PubMed]

38. Roberts, R.A.; Koehnke, J.; Besing, J. Effects of Noise and Reverberation on the Precedence Effect in Listeners with Normal Hearing and Impaired Hearing. *Am. J. Audiol.* **2003**, *12*, 96–105. [CrossRef] [PubMed]

39. Roberts, R.A.; Lister, J.J. Effects of Age and Hearing Loss on Gap Detection and the Precedence Effect: Broadband Stimuli. *J. Speech Lang. Hear. Res.* **2004**, *47*, 965–978. [CrossRef] [PubMed]

40. Bregman, A.S. Auditory Scene Analysis. In *Encyclopedia of Neuroscience*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 729–736.

41. Riecke, L.; Van Orstal, A.J.; Formisano, E. The Auditory Continuity Illusion: A Parametric Investigation and Filter Model. *Percept. Psychophys.* **2008**, *70*, 1–12. [CrossRef] [PubMed]

42. Kluender, K.R.; Jenison, R.L. Effects of Glide Slope, Noise Intensity, and Noise Duration on the Extrapolation of FM Glides through Noise. *Percept. Psychophys.* **1992**, *51*, 231–238. [CrossRef] [PubMed]

43. Kobayashi, M.; Kashino, M. Effect of Flanking Sounds on the Auditory Continuity Illusion. *PLoS ONE* **2012**, *7*, e51969. [CrossRef] [PubMed]

44. Bregman, A.S. *Auditory Scene Analysis: The Perceptual Organization of Sound*; MIT Press: Cambridge, MA, USA, 1994.

45. Petkov, C.I.; O'Connor, K.N.; Sutter, M.L. Illusory Sound Perception in Macaque Monkeys. *J. Neurosci.* **2003**, *23*, 9155–9161. [CrossRef]

46. Petkov, C.I.; O'Connor, K.N.; Sutter, M.L. Encoding of Illusory Continuity in Primary Auditory Cortex. *Neuron* **2007**, *54*, 153–165. [CrossRef]

47. Miller, C.T.; Dibble, E.; Hauser, M.D. Amodal Completion of Acoustic Signals by a Nonhuman Primate. *Nat. Neurosci.* **2001**, *4*, 783–784. [CrossRef] [PubMed]

48. Sugita, Y. Neuronal Correlates of Auditory Induction in the Cat Cortex. *Neuroreport* **1997**, *8*, 1155–1159. [CrossRef] [PubMed]

49. Cao, Q.; Parks, N.; Goldwyn, J.H. Dynamics of the Auditory Continuity Illusion. *Front. Comput. Neurosci.* **2021**, *15*, 676637. [CrossRef]

50. DeWitt, L.A.; Samuel, A.G. The Role of Knowledge-Based Expectations in Music Perception: Evidence from Musical Restoration. *J. Exp. Psychol. Gen.* **1990**, *119*, 123–144. [CrossRef] [PubMed]

51. Heinrich, A.; Carlyon, R.P.; Davis, M.H.; Johnsrude, I.S. Illusory Vowels Resulting from Perceptual Continuity: A Functional Magnetic Resonance Imaging Study. *J. Cogn. Neurosci.* **2008**, *20*, 1737–1752. [CrossRef] [PubMed]

52. Warren, R.M.; Wrightson, J.M.; Puretz, J. Illusory Continuity of Tonal and Infratonal Periodic Sounds. *J. Acoust. Soc. Am.* **1988**, *84*, 1338–1342. [CrossRef]

53. Junqua, J.C. The Lombard Reflex and Its Role on Human Listeners and Automatic Speech Recognizers. *J. Acoust. Soc. Am.* **1993**, *93*, 510–524. [CrossRef] [PubMed]

54. Derryberry, E.P.; Phillips, J.N.; Derryberry, G.E.; Blum, M.J.; Luther, D. Singing in a Silent Spring: Birds Respond to a Half-Century Soundscape Reversion during the COVID-19 Shutdown. *Science* **2020**, *370*, 575–579. [CrossRef] [PubMed]

55. Luo, J.; Hage, S.R.; Moss, C.F. The Lombard Effect: From Acoustics to Neural Mechanisms. *Trends Neurosci.* **2018**, *41*, 938–949. [CrossRef] [PubMed]

56. Heinks-Maldonado, T.H.; Houde, J.F. Compensatory Responses to Brief Perturbations of Speech Amplitude. *Acoust. Res. Lett. Online* **2005**, *6*, 131–137. [CrossRef]

57. Bauer, J.J.; Mittal, J.; Larson, C.R.; Hain, T.C. Vocal Responses to Unanticipated Perturbations in Voice Loudness Feedback: An Automatic Mechanism for Stabilizing Voice Amplitude. *J. Acoust. Soc. Am.* **2006**, *119*, 2363–2371. [CrossRef]

58. Osmanski, M.S.; Dooling, R.J. The Effect of Altered Auditory Feedback on Control of Vocal Production in Budgerigars (Melopsittacus Undulatus). *J. Acoust. Soc. Am.* **2009**, *126*, 911–919. [CrossRef]

59. Luo, J.; Kothari, N.B.; Moss, C.F. Sensorimotor Integration on a Rapid Time Scale. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6605–6610. [CrossRef]

60. Pick, H.L.; Siegel, G.M.; Fox, P.W.; Garber, S.R.; Kearney, J.K. Inhibiting the Lombard Effect. *J. Acoust. Soc. Am.* **1989**, *85*, 894–900. [CrossRef] [PubMed]

61. Brumm, H.; Todt, D. Noise-Dependent Song Amplitude Regulation in a Territorial Songbird. *Anim. Behav.* **2002**, *63*, 891–897. [CrossRef]

62. Sinnott, J.M.; Stebbins, W.C.; Moody, D.B. Regulation of Voice Amplitude by the Monkey. *J. Acoust. Soc. Am.* **1975**, *58*, 412–414. [CrossRef] [PubMed]

63. Hage, S.R.; Jiang, T.; Berquist, S.W.; Feng, J.; Metzner, W. Ambient Noise Induces Independent Shifts in Call Frequency and Amplitude within the Lombard Effect in Echolocating Bats. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 4063–4068. [CrossRef]

64. Stowe, L.M.; Golob, E.J. Evidence That the Lombard Effect Is Frequency-Specific in Humans. *J. Acoust. Soc. Am.* **2013**, *134*, 640–647. [CrossRef] [PubMed]

65. Schilling, A.; Gerum, R.; Metzner, C.; Maier, A.; Krauss, P. Intrinsic Noise Improves Speech Recognition in a Computational Model of the Auditory Pathway. *Front. Neurosci.* **2022**, *16*, 908330. [CrossRef]

66. Chatterjee, M.; Robert, M.E. Noise Enhances Modulation Sensitivity in Cochlear Implant Listeners: Stochastic Resonance in a Prosthetic Sensory System? *J. Assoc. Res. Otolaryngol.* **2001**, *2*, 159–171. [CrossRef]

67. Morse, R.P.; Holmes, S.D.; Irving, R.; McAlpine, D. Noise Helps Cochlear Implant Listeners to Categorize Vowels. *JASA Express Lett.* **2022**, *2*, 042001. [CrossRef]

68. Schaette, R.; Kempter, R. Development of Hyperactivity after Hearing Loss in a Computational Model of the Dorsal Cochlear Nucleus Depends on Neuron Response Type. *Hear. Res.* **2008**, *240*, 57–72. [CrossRef] [PubMed]

69. Kujawa, S.G.; Liberman, M.C. Adding Insult to Injury: Cochlear Nerve Degeneration after "Temporary" Noise-Induced Hearing Loss. *J. Neurosci.* **2009**, *29*, 14077–14085. [CrossRef] [PubMed]

70. Schaette, R.; McAlpine, D. Tinnitus with a Normal Audiogram: Physiological Evidence for Hidden Hearing Loss and Computational Model. *J. Neurosci.* **2011**, *31*, 13452–13457. [CrossRef] [PubMed]

71. Monaghan, J.J.M.; Garcia-Lazaro, J.A.; McAlpine, D.; Schaette, R. Hidden Hearing Loss Impacts the Neural Representation of Speech in Background Noise. *Curr. Biol.* **2020**, *30*, 4710–4721.e4. [CrossRef]

72. Bakay, W.M.H.; Anderson, L.A.; Garcia-Lazaro, J.A.; McAlpine, D.; Schaette, R. Hidden Hearing Loss Selectively Impairs Neural Adaptation to Loud Sound Environments. *Nat. Commun.* **2018**, *9*, 4298. [CrossRef] [PubMed]

73. Schilling, A.; Sedley, W.; Gerum, R.; Metzner, C.; Tziridis, K.; Maier, A.; Schulze, H.; Zeng, F.-G.; Friston, K.J.; Krauss, P. Predictive Coding and Stochastic Resonance as Fundamental Principles of Auditory Phantom Perception. *Brain* **2023**, *146*, 4809–4825. [CrossRef]

74. Sohoglu, E.; Chait, M. Detecting and Representing Predictable Structure during Auditory Scene Analysis. *eLife* **2016**, *5*, e19113. [CrossRef] [PubMed]

75. Aman, L.; Picken, S.; Andreou, L.-V.; Chait, M. Sensitivity to Temporal Structure Facilitates Perceptual Analysis of Complex Auditory Scenes. *Hear. Res.* **2021**, *400*, 108111. [CrossRef]

76. Chait, M.; Ruff, C.C.; Griffiths, T.D.; McAlpine, D. Cortical Responses to Changes in Acoustic Regularity Are Differentially Modulated by Attentional Load. *Neuroimage* **2012**, *59*, 1932–1941. [CrossRef] [PubMed]

77. McDermott, J.H.; Schemitsch, M.; Simoncelli, E.P. Summary Statistics in Auditory Perception. *Nat. Neurosci.* **2013**, *16*, 493–498. [CrossRef]

78. Agus, T.R.; Thorpe, S.J.; Pressnitzer, D. Rapid Formation of Robust Auditory Memories: Insights from Noise. *Neuron* **2010**, *66*, 610–618. [CrossRef] [PubMed]

79. Blom, J.D. Chapter 24—Auditory Hallucinations. In *Handbook of Clinical Neurology*; Aminoff, M.J., Boller, F., Swaab, D.F., Eds.; Elsevier: Amsterdam, The Netherlands, 2015; Volume 129, pp. 433–455.

80. Culling, J.F.; Summerfield, A.Q.; Marshall, D.H. Dichotic Pitches as Illusions of Binaural Unmasking. I. Huggins' Pitch and the "Binaural Edge Pitch". *J. Acoust. Soc. Am.* **1998**, *103*, 3509–3526. [CrossRef] [PubMed]

81. Blauert, J.; Lindemann, W. Spatial Mapping of Intracranial Auditory Events for Various Degrees of Interaural Coherence. *J. Acoust. Soc. Am.* **1986**, *79*, 806–813. [CrossRef] [PubMed]

82. Hall, D.A.; Barrett, D.J.K.; Akeroyd, M.A.; Summerfield, A.Q. Cortical Representations of Temporal Structure in Sound. *J. Neurophysiol.* **2005**, *94*, 3181–3191. [CrossRef] [PubMed]

83. Akeroyd, M.A.; Moore, B.C.; Moore, G.A. Melody Recognition Using Three Types of Dichotic-Pitch Stimulus. *J. Acoust. Soc. Am.* **2001**, *110 Pt 1*, 1498–1504. [CrossRef]

84. Freiwald, W.; Tsao, D. *Taking Apart the Neural Machinery of Face Processing*; Oxford University Press: Oxford, UK, 2011.

85. Mesgarani, N.; David, S.V.; Fritz, J.B.; Shamma, S.A. Mechanisms of Noise Robust Representation of Speech in Primary Auditory Cortex. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6792–6797. [CrossRef] [PubMed]

86. Moore, R.C.; Channing Moore, R.; Lee, T.; Theunissen, F.E. Noise-Invariant Neurons in the Avian Auditory Cortex: Hearing the Song in Noise. *PLoS Comput. Biol.* **2013**, *9*, e1002942. [CrossRef]

87. Schneider, D.M.; Woolley, S.M.N. Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes. *Neuron* **2013**, *79*, 141–152. [CrossRef] [PubMed]

88. Rabinowitz, N.C.; Willmore, B.D.B.; King, A.J.; Schnupp, J.W.H. Constructing Noise-Invariant Representations of Sound in the Auditory Pathway. *PLoS Biol.* **2013**, *11*, e1001710. [CrossRef]

89. Kell, A.J.E.; McDermott, J.H. Invariance to Background Noise as a Signature of Non-Primary Auditory Cortex. *Nat. Commun.* **2019**, *10*, 3958. [CrossRef]

90. Bar-Yosef, O.; Nelken, I. The Effects of Background Noise on the Neural Responses to Natural Sounds in Cat Primary Auditory Cortex. *Front. Comput. Neurosci.* **2007**, *1*, 3. [CrossRef]

91. Wiegrebe, L.; Winter, I.M. Temporal Representation of Iterated Rippled Noise as a Function of Delay and Sound Level in the Ventral Cochlear Nucleus. *J. Neurophysiol.* **2001**, *85*, 1206–1219. [CrossRef] [PubMed]

92. Woolley, S.M.N.; Casseday, J.H. Processing of Modulated Sounds in the Zebra Finch Auditory Midbrain: Responses to Noise, Frequency Sweeps, and Sinusoidal Amplitude Modulations. *J. Neurophysiol.* **2005**, *94*, 1143–1157. [CrossRef] [PubMed]

93. Khatami, F.; Escabí, M.A. Spiking Network Optimized for Word Recognition in Noise Predicts Auditory System Hierarchy. *PLoS Comput. Biol.* **2020**, *16*, e1007558. [CrossRef]

94. Souffi, S.; Nodal, F.R.; Bajo, V.M.; Edeline, J.-M. When and How Does the Auditory Cortex Influence Subcortical Auditory Structures? New Insights About the Roles of Descending Cortical Projections. *Front. Neurosci.* **2021**, *15*, 690223. [CrossRef]

95. Souffi, S.; Lorenzi, C.; Varnet, L.; Huetz, C.; Edeline, J.-M. Noise-Sensitive But More Precise Subcortical Representations Coexist with Robust Cortical Encoding of Natural Vocalizations. *J. Neurosci.* **2020**, *40*, 5228–5246. [CrossRef] [PubMed]

96. Robinson, B.L.; Harper, N.S.; McAlpine, D. Meta-Adaptation in the Auditory Midbrain under Cortical Influence. *Nat. Commun.* **2016**, *7*, 13442. [CrossRef] [PubMed]

97. Norman-Haignere, S.V.; McDermott, J.H. Neural Responses to Natural and Model-Matched Stimuli Reveal Distinct Computations in Primary and Nonprimary Auditory Cortex. *PLoS Biol.* **2018**, *16*, e2005127. [CrossRef]

98. Selezneva, E.; Gorkin, A.; Budinger, E.; Brosch, M. Neuronal Correlates of Auditory Streaming in the Auditory Cortex of Behaving Monkeys. *Eur. J. Neurosci.* **2018**, *48*, 3234–3245. [CrossRef]

99. Cruces-Solís, H.; Jing, Z.; Babaev, O.; Rubin, J.; Gür, B.; Krueger-Burg, D.; Strenzke, N.; de Hoz, L. Auditory Midbrain Coding of Statistical Learning That Results from Discontinuous Sensory Stimulation. *PLoS Biol.* **2018**, *16*, e2005114. [CrossRef]

100. Chen, C.; Cruces-Solís, H.; Ertman, A.; de Hoz, L. Subcortical Coding of Predictable and Unsupervised Sound-Context Associations. *Curr. Res. Neurobiol.* **2023**, *5*, 100110. [CrossRef] [PubMed]

101. Asokan, M.M.; Williamson, R.S.; Hancock, K.E.; Polley, D.B. Inverted Central Auditory Hierarchies for Encoding Local Intervals and Global Temporal Patterns. *Curr. Biol.* **2021**, *31*, 1762–1770.e4. [CrossRef]

102. Henin, S.; Turk-Browne, N.B.; Friedman, D.; Liu, A.; Dugan, P.; Flinker, A.; Doyle, W.; Devinsky, O.; Melloni, L. Learning Hierarchical Sequence Representations across Human Cortex and Hippocampus. *Sci. Adv.* **2021**, *7*, eabc4530. [CrossRef]

103. Happel, M.F.K.; Hechavarria, J.C.; de Hoz, L. Editorial: Cortical-Subcortical Loops in Sensory Processing. *Front. Neural Circuits* **2022**, *16*, 851612. [CrossRef]

104. de Hoz, L.; Busse, L.; Hechavarria, J.C.; Groh, A.; Rothermel, M. SPP2411: "Sensing LOOPS: Cortico-Subcortical Interactions for Adaptive Sensing". *Neuroforum* **2022**, *28*, 249–251. [CrossRef]

105. Cardin, J.A. Functional Flexibility in Cortical Circuits. *Curr. Opin. Neurobiol.* **2019**, *58*, 175–180. [CrossRef] [PubMed]

106. Saldaña, E. All the Way from the Cortex: A Review of Auditory Corticosubcollicular Pathways. *Cerebellum* **2015**, *14*, 584–596. [CrossRef] [PubMed]

107. Nakamoto, K.T.; Mellott, J.G.; Killius, J.; Storey-Workley, M.E.; Sowick, C.S.; Schofield, B.R. Ultrastructural Examination of the Corticollicular Pathway in the Guinea Pig: A Study Using Electron Microscopy, Neural Tracers, and GABA Immunocytochemistry. *Front. Neuroanat.* **2013**, *7*, 13. [CrossRef] [PubMed]

108. Yudintsev, G.; Asilador, A.R.; Sons, S.; Vaithiyalingam Chandra Sekaran, N.; Coppinger, M.; Nair, K.; Prasad, M.; Xiao, G.; Ibrahim, B.A.; Shinagawa, Y.; et al. Evidence for Layer-Specific Connectional Heterogeneity in the Mouse Auditory Corticocollicular System. *J. Neurosci.* **2021**, *41*, 9906–9918. [CrossRef] [PubMed]

109. Bajo, V.M.; Nodal, F.R.; Moore, D.R.; King, A.J. The Descending Corticocollicular Pathway Mediates Learning-Induced Auditory Plasticity. *Nat. Neurosci.* **2010**, *13*, 253–260. [CrossRef]

110. Bajo, V.M.; King, A.J. Cortical Modulation of Auditory Processing in the Midbrain. *Front. Neural Circuits* **2012**, *6*, 114. [CrossRef] [PubMed]

111. Gao, E.; Suga, N. Experience-Dependent Plasticity in the Auditory Cortex and the Inferior Colliculus of Bats: Role of the Corticofugal System. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8081–8086. [CrossRef]

112. Bajo, V.M.; Nodal, F.R.; Korn, C.; Constantinescu, A.O.; Mann, E.O.; Boyden, E.S., 3rd; King, A.J. Silencing Cortical Activity during Sound-Localization Training Impairs Auditory Perceptual Learning. *Nat. Commun.* **2019**, *10*, 3075. [CrossRef] [PubMed]

113. Xiong, X.R.; Liang, F.; Zingg, B.; Ji, X.-Y.; Ibrahim, L.A.; Tao, H.W.; Zhang, L.I. Auditory Cortex Controls Sound-Driven Innate Defense Behaviour through Corticofugal Projections to Inferior Colliculus. *Nat. Commun.* **2015**, *6*, 7224. [CrossRef]

114. Dean, I.; Robinson, B.L.; Harper, N.S.; McAlpine, D. Rapid Neural Adaptation to Sound Level Statistics. *J. Neurosci.* **2008**, *28*, 6430–6438. [CrossRef] [PubMed]

115. Wen, B.; Wang, G.I.; Dean, I.; Delgutte, B. Time Course of Dynamic Range Adaptation in the Auditory Nerve. *J. Neurophysiol.* **2012**, *108*, 69–82. [CrossRef]

116. Dean, I.; Harper, N.S.; McAlpine, D. Neural Population Coding of Sound Level Adapts to Stimulus Statistics. *Nat. Neurosci.* **2005**, *8*, 1684–1689. [CrossRef]

117. Watkins, P.V.; Barbour, D.L. Level-Tuned Neurons in Primary Auditory Cortex Adapt Differently to Loud versus Soft Sounds. *Cereb. Cortex* **2011**, *21*, 178–190. [CrossRef]

118. Willmore, B.D.B.; Schoppe, O.; King, A.J.; Schnupp, J.W.H.; Harper, N.S. Incorporating Midbrain Adaptation to Mean Sound Level Improves Models of Auditory Cortical Processing. *J. Neurosci.* **2016**, *36*, 280–289. [CrossRef]