



Top-Down Inference in the Auditory System: Potential Roles for Corticofugal Projections

Alexander Asilador^{1,2} and Daniel A. Llano^{1,2,3*}

¹Neuroscience Program, The University of Illinois at Urbana-Champaign, Champaign, IL, United States, ²Beckman Institute for Advanced Science and Technology, Urbana, IL, United States, ³Molecular and Integrative Physiology, The University of Illinois at Urbana-Champaign, Champaign, IL, United States

It has become widely accepted that humans use contextual information to infer the meaning of ambiguous acoustic signals. In speech, for example, high-level semantic, syntactic, or lexical information shape our understanding of a phoneme buried in noise. Most current theories to explain this phenomenon rely on hierarchical predictive coding models involving a set of Bayesian priors emanating from high-level brain regions (e.g., prefrontal cortex) that are used to influence processing at lower-levels of the cortical sensory hierarchy (e.g., auditory cortex). As such, virtually all proposed models to explain top-down facilitation are focused on intracortical connections, and consequently, subcortical nuclei have scarcely been discussed in this context. However, subcortical auditory nuclei receive massive, heterogeneous, and cascading descending projections at every level of the sensory hierarchy, and activation of these systems has been shown to improve speech recognition. It is not yet clear whether or how top-down modulation to resolve ambiguous sounds calls upon these corticofugal projections. Here, we review the literature on top-down modulation in the auditory system, primarily focused on humans and cortical imaging/recording methods, and attempt to relate these findings to a growing animal literature, which has primarily been focused on corticofugal projections. We argue that corticofugal pathways contain the requisite circuitry to implement predictive coding mechanisms to facilitate perception of complex sounds and that top-down modulation at early (i.e., subcortical) stages of processing complement modulation at later (i.e., cortical) stages of processing. Finally, we suggest experimental approaches for future studies on this topic.

OPEN ACCESS

Edited by:

Julio C. Hechavarría,
Goethe University Frankfurt, Germany

Reviewed by:

Kirill Vadimovich Nourski,
The University of Iowa, United States
Kasia M. Bieszczad,
The State University of New Jersey,
United States

*Correspondence:

Daniel A. Llano
d-llano@illinois.edu

Keywords: auditory, cortex, thalamus, colliculus, top-down, speech perception, descending, medial geniculate body

Received: 08 October 2020

Accepted: 17 December 2020

Published: 22 January 2021

Citation:

Asilador A and Llano DA
(2021) Top-Down Inference in the
Auditory System: Potential Roles for
Corticofugal Projections.
Front. Neural Circuits 14:615259.
doi: 10.3389/fncir.2020.615259

INTRODUCTION

We effortlessly navigate a world filled with complex sounds. Despite challenging listening environments, such as having a conversation on a windy day, talking over a poor cell phone connection, or presenting a poster at a busy scientific meeting, the auditory system routinely extracts the meaning of signals corrupted by noise. One type of cue that may be used to perform this operation is the linguistic or acoustic context within which a sound exists. For example, it has long been known that high-level information about the nature of ambiguous speech sounds

can dramatically enhance the ability to recognize these sounds (Miller et al., 1951; O'Neill, 1957; and reviewed in Davis and Johnsrude, 2007; Obleser, 2014). Also, acoustic perception and peripheral auditory responses in humans are strongly influenced by preceding non-speech acoustic stimuli (Lotto and Kluender, 1998; Skoe and Kraus, 2010), suggesting that contextual cueing may be a general mechanism used by the auditory system to deal with ambiguity. Contextual cueing is also of clinical importance as many individuals with language-related disorders, such as aphasia, autism, auditory processing disorder, and dyslexia, have difficulties using high-level contextual cues to disambiguate noisy or degraded sound stimuli (Tseng et al., 1993; Grindrod and Baum, 2002; Fink et al., 2006; Stewart and Ota, 2008; Chandrasekaran et al., 2009; Moore D. R., 2012).

The process of using prior knowledge to influence the processing of sensory information is referred to as “top-down modulation.” Originally described as “unconscious influence” by Helmholtz in the 1800s (Von Helmholtz, 1867), top-down modulation is a ubiquitous process that is seen across all sensory systems (Kobayashi et al., 2004; Haegens et al., 2011; Andersson et al., 2018). It is believed that the major roles of top-down modulation are to select certain sensory features over others in a cluttered sensory environment to favor encoding information that is more meaningful for the organism. On the latter point, meaningful information is often defined by the statistical regularity with which those features are encountered in the environment, a key point exploited by most of the experimental paradigms involving repetitive stimulation of a particular region of cortex (e.g., Gao and Suga, 2000; Yan and Ehret, 2002).

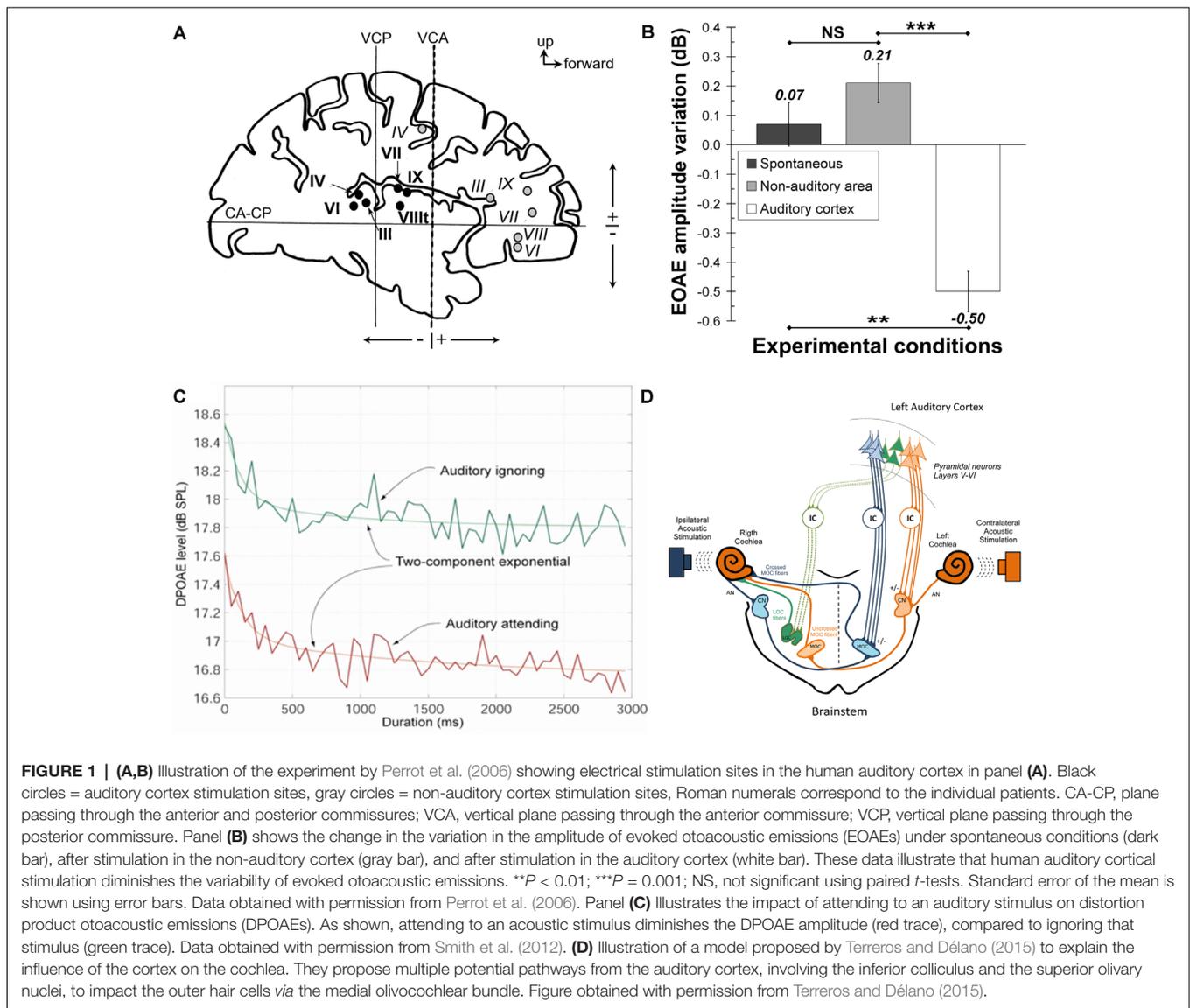
The neural substrates for top-down modulation are not well understood. Sensory systems are hierarchically organized such that sensory information ascends through a series of brain regions before reaching the primary sensory cortex (e.g., the primary auditory cortex). Canonically, the primary sensory cortex sends projections to secondary sensory cortical areas, which then project to areas outside of the sensory pathway, typically including areas of the prefrontal cortex. Also, virtually all of these “ascending” connections are associated with a returning “descending” connection, which in some cases contain axons that greatly outnumber the corresponding ascending connection. In some cases, the descending connections “skip” levels and send projections to areas that do not have a direct corresponding ascending connection (e.g., the projection from the cortex to the tectum or to the corpus striatum). Virtually all current models that describe the use of top-down modulation to facilitate auditory processing have focused on intracortical projections [e.g., from the frontal cortex to auditory cortex or from secondary auditory cortical fields to the primary auditory cortex (Zekveld et al., 2006; Hannemann et al., 2007; Sohoglu et al., 2012; Chennu et al., 2013, 2016; Hofmann-Shen et al., 2020)]. What is often left out of the discussion, however, are the massive and heterogeneous projections emanating from the auditory cortex that target virtually every level of the subcortical auditory system (herein “corticofugal projections”) and, through cascading projections, impacting the most peripheral component: the cochlea

(Xiao and Suga, 2002; León et al., 2012; Dragicevic et al., 2015; Jäger and Kössl, 2016).

This focus on cortical mechanisms of top-down modulation has existed despite the data demonstrating that descending influences can alter primary auditory input through the cochlear efferent system. For example, attentional tasks and prior linguistic knowledge modulate efferent projections to the cochlea (Collet et al., 1994; Marian et al., 2018), electrical stimulation of the human auditory cortex modulates cochlear activity (Perrot et al., 2006), and activation of subcortical auditory pathways to the cochlea facilitate speech recognition in challenging listening situations (De Boer and Thornton, 2008; Smith et al., 2012; Srinivasan et al., 2012; Mishra and Lutman, 2014; Shastri et al., 2014). As shown in **Figures 1A,B**, electrical stimulation of the human auditory cortex (but not non-auditory cortex) diminishes the mean amplitude and the variation in the amplitude of evoked otoacoustic emissions. Also, auditory attention leads to a decline in the amplitude of otoacoustic emissions, which are generated by the cochlea (**Figure 1C**). The projections from the auditory cortex that lead to modulation of the cochlea have been reviewed by Terreros and Déano (2015). They proposed a cascading model of multiple parallel pathways connecting the auditory cortex, inferior colliculus, cochlear nucleus, and superior olivary nuclei (including a direct projection from the auditory cortex to neurons making up the medial olivocochlear pathway; Mulders and Robertson, 2000) as potential neural substrates for these findings (**Figure 1D**; Terreros and Déano, 2015). Here, we attempt to link the bodies of literature on intracortical top-down modulation for processing of complex sounds (which has primarily been done in humans, with some notable exceptions; García-Rosales et al., 2020; Yin et al., 2020) and corticofugal modulation of subcortical auditory processing regions (which has primarily been done in animals), to develop a better understanding of the potential role of corticofugal projections in the disambiguation of corrupted acoustic signals.

EVIDENCE FOR TOP-DOWN MODULATION IN THE AUDITORY SYSTEM: HUMAN STUDIES

When engaged with acoustic stimuli, the goal at the behavioral level is the coherent perception of an object in its environment. In the auditory system, one of the earliest models used to describe perception was auditory scene analysis. The term was coined by Albert Bregman, a psychologist at McGill University (Bregman, 1994). He explored the idea that elements of a sound stimulus are grouped by the similarity of the components of a sound. These bottom-up features include the pitch, harmonicity, rhythmicity, similarity of sound, and timing of the sounds. Research in perceptual computing has shown some success in forming the foundation of scene analysis, where the computational model is capable of object detection, component extraction, and separation of sources in real-world situations (Smaragdis, 2001). However, when the level of ambiguity increases, object separation becomes much more difficult. Researchers have



investigated the effect of attention to resolve ambiguities, such as the separation of objects from distractors and noisy environments. For example, van Noorden (1971) examined stream segregation by presenting pure tones, tone A and tone B, to listeners. The stimulus was presented as a sequence of alternating A and B tones, but every second B tone was omitted. The two tones differed by a pitch for each experiment, and this difference was distinguished as either a denoted “small,” “intermediate,” or “large” difference. For small differences, the tones were perceived as a single rhythm and result in the perceptual fusion of the two tones. For large differences, the resulting perceived sound led to a separation of the two sounds, where the A tone was presented twice as fast as the B tone. For intermediate differences, the listeners either perceived either a fusion or fission of the two sounds based on the subject, however, the subjects can influence what type they hear based on the instructions given to the subjects. Thus, attentional bias

can determine the nature of a percept when ambiguous signals are presented.

The effects of top-down modulation on bottom-up processing are particularly notable during speech perception. Any given speech unit is not represented solely by the instantaneous components of sound (frequency content and intensity) but is a time-varying cognitive construct whereby a combination of phonemes or acoustical patterns are used to represent a unit of speech. The same speech sounds vary from speaker to speaker and speech sounds may change based on their preceding or following sounds (coarticulation; Moore B. C., 2012). Yet, listeners can understand phrases and dialogue from different speakers without difficulty. As outlined by Davis and Johnsruide (2007), this form of perception is experience-driven and is demonstrated from an analysis by Fodor and Bever (1965) on the inclusion of clicks in a speech, as seen in speakers of Sub-Saharan languages. Such

psychoacoustic tests have revealed that the clicks are not perceptually heard in individuals who have not acquired this language. The argument here is that speech understanding is a perceptual process such that humans cognitively reorganize the acoustic input stream based on our experience with acoustic stimuli.

Several core perceptual processes are needed to effectuate speech perception in the face of widely varying sensory stimuli. One is categorical perception—the tendency to perceive acoustic stimuli as belonging to distinct categories despite having their stimulus properties vary on a continuum (e.g., perceiving a phoneme as either voiced vs. unvoiced despite having a gradual change in voice onset time). Another perceptual process that is key to understanding corrupted speech is a perceptual fill-in. In speech, this is typically referred to as the “phonemic restoration effect” (Warren, 1970) and describes the process of perceptually filling in noise-filled gaps in speech with the missing phoneme, analogous to filling in the contour of a partially obscured or partially-constructed visual object (e.g., Kanizsa objects). A third core perceptual process needed for speech processing is segmentation. That is, knowing the start and the stop of a meaningful acoustic signal. Generally, speech does not provide clear temporal demarcations between meaningful utterances, and these have to be inferred by the listener. Finally, stream segregation—the ability to perceptually separate different auditory objects whose waveforms are intermingled—is key to deciphering speech buried in noise. Although these core perceptual processes for speech understanding can potentially be explained solely *via* bottom-up processes (see Norris et al., 2000 for arguments in favor of a purely bottom-up approach to speech processing), as will be reviewed below, they are all strongly influenced by top-down factors.

Early evidence that lexical or semantic context could be used to facilitate the categorical perception of speech in noise was provided by Miller et al. (1951). They reported that the intelligibility of a word is enhanced when the appropriate context is provided. For example, the word “trees” buried in noise is more intelligible if it is preceded by the phrase “Apples grow on ____”. Later work established that this effect is present at the lexical level (Ganong effect) such that preceding phonemes could increase the intelligibility of subsequent phonemes in words compared to nonwords (e.g., “task” vs. “dask”; Ganong, 1980).

Non-auditory cues can also be used to facilitate categorical perception. For example, observing the mouth movements of a speaker or seeing a written representation of a word before the obscured sound both facilitated perceptual performance (Sohoglu et al., 2012, 2014; Getz and Toscano, 2019; Pinto et al., 2019). For example, providing a written example of a semantically-associated word (e.g., “MASHED”) before an acoustic representation of a word with ambiguous voice onset time (e.g., “potatoes”), facilitated the categorical perception of the initial consonant more than unrelated visual primes (Getz and Toscano, 2019). This use of cross-modal semantic priming modulated the earliest electroencephalography (EEG) peak examined by the investigators, the N1 peak, thought to be related to primary auditory cortex activation

(Hillyard et al., 1973; Näätänen and Picton, 1987). Also, to compare the contributions of frontal vs. temporal cortex in a similar task, Sohoglu et al. found that use of written word prior information to disambiguate a vocoded speech sound was associated with inferior frontal gyrus activation using a combined EEG/magnetoencephalography (MEG) approach. In contrast, manipulations of the number of frequency bands available (thus increasing the bottom-up detail in the stimulus), activated auditory areas of the superior temporal gyrus (Sohoglu et al., 2012). These data are in line with a fronto-temporal hypothesis about descending control (Tzourio et al., 1997; Braga et al., 2013; Cope et al., 2017).

Concerning perceptual fill-in, the influence of context on phonemic restoration has been extensively examined, even from the earliest descriptions of the restoration phenomenon. For example, Marslen-Wilson demonstrated in 1975 that phonemic restoration was much more common when the target word was placed in the appropriate semantic and grammatical context and that the third syllable of a word was much more likely to be restored than the first syllable (Marslen-Wilson, 1975), suggesting that within-word context is an important cue. Expectation effects were also found by Samuel in 1981 who showed that words with a syllable replaced by noise were more likely to be reported as intact words if those words were incorporated into a sentence (Samuel, 1981). Samuel later (Samuel, 1997) showed that phonemic restoration introduced adaptation effects similar to those predicted by previous top-down models (e.g., the TRACE model; McClelland and Elman, 1986). More recently, it has been shown that the phonemic restoration effect remains intact despite voice discontinuities pre-and post-noise gap. That is, listeners were able to perceptually fill-in the gap despite the absence of spectral overlap between the pre-and post-gap voice, suggesting that other cues, such as linguistic context, are driving the filling-in phenomenon (McGettigan et al., 2013).

The third core perceptual process needed to disambiguate noisy speech is segmentation. Because most languages do not have clear acoustic demarcations separating meaningful utterances in speech, segmentation between words and sentences must be inferred (e.g., “mother’s cold” vs. “mother scold”), and thus represents a key component of top-down speech perception (Davis and Johnsruide, 2007). Indeed a common complaint among most learners of a new language is not knowing where words start and end. Multiple potential cues can assist in this segmentation, such as loudness (stresses on particular syllables), word knowledge, semantic context, etc. Mattys et al. found that when multiple conflicting cues were available, listeners relied on higher-level cues (e.g., sentence context) rather than lower-level cues (e.g., word stress). They proposed a hierarchical organization with lexical knowledge occupying the highest level and what they referred to as “metrical prosody” (syllable stresses) at the bottom (Mattys et al., 2005). Supporting the idea that word knowledge plays a role in lexical segmentation is the finding by Cunillera et al. (2010) that knowing a small number of “anchor” words in a novel language facilitated the ability

to appropriately segment that language into meaningful units. Similar knowledge-based facilitation of segmentation of musical phrases has been observed, suggesting that top-down facilitation of segmentation may be a general property of the auditory system (Silva et al., 2019).

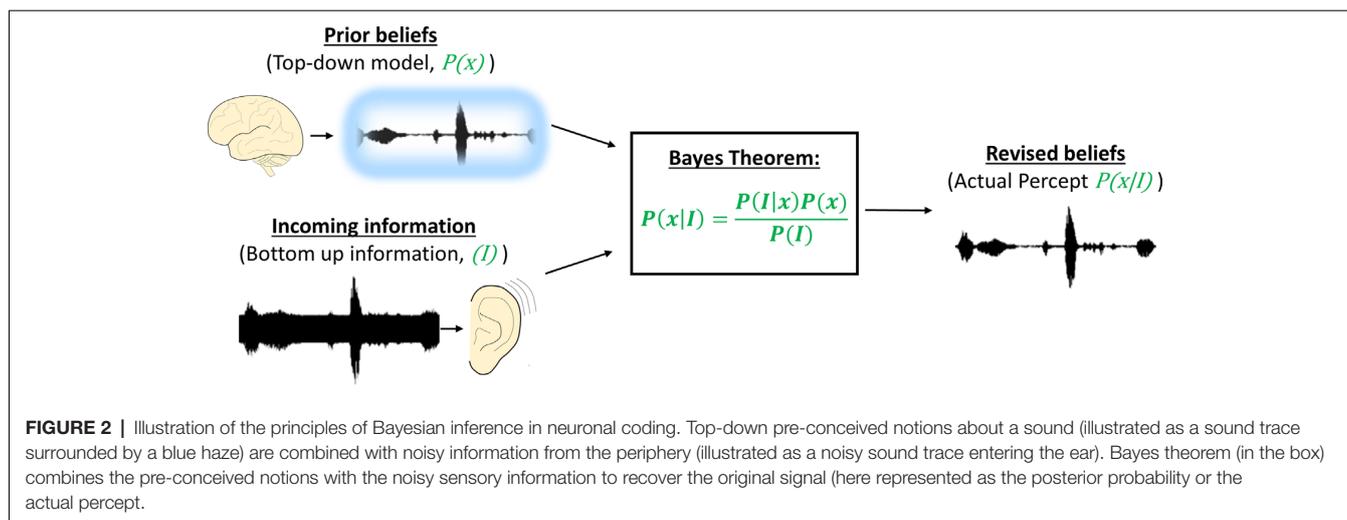
Another key requirement for inferring speech content under noisy conditions is the ability to separate competing sound streams. This process is multifaceted and involves both bottom-up cues (e.g., different pitch contours or spatial locations of different sources, as described above; Bregman, 1994) and top-down cues. Many investigators have established that bottom-up cues are sufficient to separate sound sources (often referred to as “sound streams”) when the physical characteristics of the sound sources are distinct (Scholes et al., 2015). However, when there is substantial overlap between them, as is often the case in a sound-cluttered real-world environment, top-down cues become critical. Several studies have been done using such cluttered stimuli and have presented a priming stimulus containing the target and have observed a marked improvement in identifying the target (Freyman et al., 2004; Jones and Freyman, 2012; Wang et al., 2019). For example, Wang et al. examined the ability to separate two simultaneously-presented spectrally- and temporally- overlapping talkers without spatial cues. The presence of the target sound played before the simultaneously-presented sounds greatly facilitated the recognition of the target. This recognition was also associated with increased phase-locking of the superior temporal gyrus and sulcus MEG signals to the speech envelope (Wang et al., 2019). These data suggest that in the absence of bottom-up cues to separate sound sources, knowledge-based cues can be used and that this knowledge modulates processing in areas of the auditory cortex.

A common class of paradigms to study the various perceptual processes involved in auditory top-down modulation in humans is the oddball or omission paradigm. Such paradigms typically involve repetition of a particular sound, followed by an “oddball” (e.g., AAAAB), or the absence of sound (e.g., AAAA_). This paradigm or variations of it (e.g., presenting a global deviant such as AAAAA in the setting of a long series of AAAAB stimuli) have been heavily employed in the neuroscience literature. Oddballs typically evoke a voltage change measured at the scalp known as the mismatch negativity (MMN). The presence of the mismatch negativity has been taken as evidence of a core component of predictive coding—prediction error—and has thus been promoted as evidence for top-down modulation in the auditory system. The mapping of MMN onto top-down processing mechanisms is still not clear. The presence of some forms of MMN (sensitivity to local, rather than global deviants) in sleep or under anesthesia (Loewy et al., 2000; Nourski et al., 2018), which would be inconsistent with an active inferential process, suggests that bottom-up effects (such as habituation to repeated stimuli) may play a role. More modern instantiations of the oddball paradigm comparing responses to local vs. global deviants have shown that global deviants may be more vulnerable to anesthesia (Nourski et al., 2018), suggesting that this form of predictive error may better reflect active top-down control mechanisms.

COMPUTATIONAL PRINCIPLES OF TOP-DOWN MODULATION

Various models have been proposed to understand how contextual cues can influence sensory processing. Predictive coding is a general framework by which context, in the form of predictions about incoming data, can shape the properties of sensory-responsive neurons. Early instantiations of predictive coding algorithms were primarily focused on increasing the efficiency of the coding because of predictive coding’s ability to reduce redundancies in data streams (Srinivasan et al., 1982), similar to bandwidth compression required to transmit large images. Notably, efficiency in terms of the number of neuronal connections does not appear to be a design principle of descending systems in the brain. These systems are massive and typically dwarf ascending projections, so it seems unlikely that they evolved to maximize the efficiency of coding in lower centers. It is more likely that these large, presumably energy-expensive systems, evolved to increase the accuracy of identifying causes of sensory inputs. To this end, approaches that have been shown to increase the accuracy of sensory estimation, such as Bayesian estimation, have been postulated to be of use. Such schemes involve the generation of a prediction about the outside world (a Bayesian prior) that, when combined with noisy or degraded sensory information, leads to an optimal estimate of the cause of the sensory signal (the posterior probability), see **Figure 2**. The Bayesian priors are based on previous experience with the world and thus are updated by experience. Several studies have shown that in the setting of sensory uncertainties, humans combine contextual information and sensory information in Bayes-optimal ways (Jacobs, 1999; Ernst and Banks, 2002; Battaglia et al., 2003). More general models that attempt to explain neural processing based on similar principles (e.g., the Free Energy Principle) have been proposed (Friston and Kiebel, 2009). In practice, most predictive coding models involve a prediction, which is compared to sensory input. When the two are unmatched, a “prediction error” occurs (increasing the Free Energy), which is used as a learning signal to modify the internal model. This scheme is consistent with the large body of work showing enhanced neural responses to unpredicted stimuli (e.g., the MMN, reviewed above). However, as described in “Neural Models of Top-Down Modulation,” section this model has challenges both at the neural implementation level and at the level of linking neural responses to behavior.

Precisely how to link internal models with incoming information has been an open question. In modeling studies, the integration of predictive cues with incoming information has been implemented using several approaches. One approach has adapted linear systems theory and estimation theory into a model of the visual system. Rao and Ballard (1997) condensed the complexity of the visual system into a series of calculations that are inspired by work in minimum mean squared error estimation (MMSE): the Kalman filter (Kalman, 1960). The goal of MMSE is to estimate the internal (unknown) state of a system based on observation of noisy sensors to predict the next state. The Kalman filter is a linear estimator that assumes



the noise from the environment is Gaussian. Further, any noise imparted by the internal state itself is pairwise uncorrelated to the noise of the sensor. This filter was used in an early model of hierarchical predictive coding in the visual system that, when trained on natural images, recapitulated some of the receptive field properties of early visual cortical neurons (Rao and Ballard, 1997). The model itself applies an extended form of the Kalman filter, capable of learning and prediction, with the learning rule obtained by the expectation-maximization (EM) algorithm formulated to mimic Hebbian learning. It can be shown that under Gaussian conditions that the Kalman filter is equivalent to the Bayes filter (Chen, 2003). A similar model to the extended Kalman filter initially proposed by Rao and Ballard (1997) has implemented a generative dynamical system in place of the Kalman filter to capture nonlinearities of neural activation, and a learning scheme that takes into account the extra-classical effects experimentally observed in the visual system (Rao and Ballard, 1999).

Under non-Gaussian conditions, a more general implementation approach that has been used is the particle filter, as proposed by Lee and Mumford (2003). The calculations between the Kalman filter and particle filter are not similar, as the particle filter generates the likelihood weighting of states from the input and previous weights, followed by resampling of the input. While the Kalman filter is expressed by a linear operation, particle filters are constructed similarly to Markov chains to estimate the state of a given observation. The difference is that the dimensionality of the model is reduced by only looking at a weighted probability of being at a state instead of the total probability. This requires sampling a portion of the complete observation and estimating the weighted probability of the object being at some state. Subsequent re-sampling is performed with the weighted probabilities fed back into the model to more confidently estimate the state. In Lee and Mumford's influential 2003 article, the authors introduce a concept of a particle filter-based model that hypothesizes that cortical connections are responsible for the calculations but interact in a way where each neuron represents specific events in the external

world (i.e., features of an object; Lee and Mumford, 2003). It is described as a generative model, calculating the likelihood of the state hierarchically. Single neuron activation indicates a specific event in the external environment. The external environment shows the co-activation of specific patterns, and the state of the hidden variable depends on the state at the previous time step. Synchronized activity in a population of neurons contributes to the image. Here, the activity of superficial pyramidal cells correspond to the bottom-up messages, and the deeper pyramidal cells reflect top-down messages. Current state is conditionally independent of other past states.

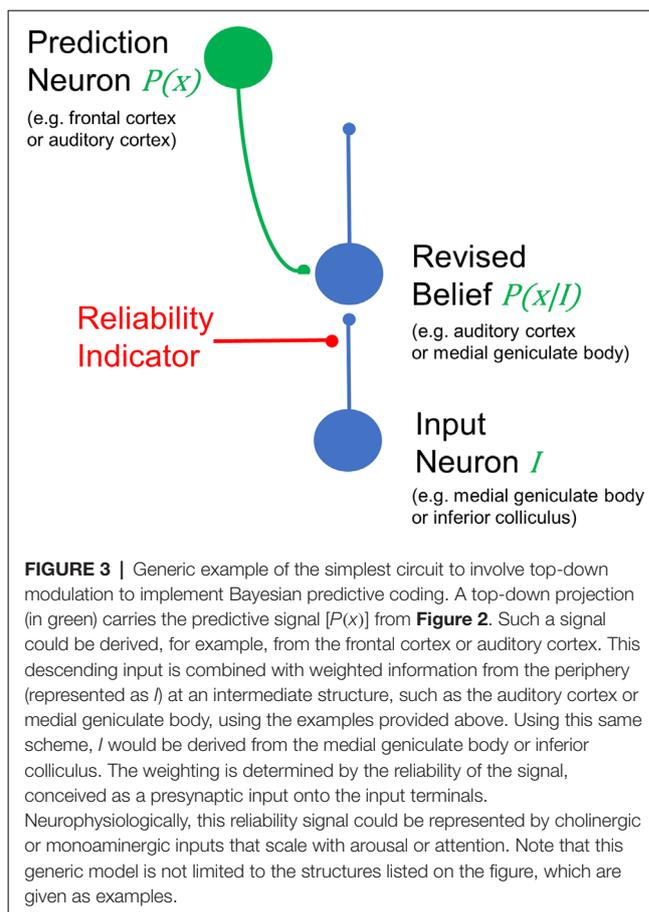
NEURAL MODELS OF TOP-DOWN MODULATION

Neural models employed in predictive coding algorithms have relied heavily on descending connections between cortical areas. For example, in an early large-scale iterative model of visual cortico-cortical interactions that implemented predictive coding, a hierarchical network was proposed, with the lowest level focusing on a small portion of the image (local image patches) at a short time scale, and each subsequent level in the hierarchy representing increasing feature complexity such as larger spatial and time scales (Rao and Ballard, 1999). In this model, it was argued that each level in the hierarchy first starts at the inputs from the visual thalamus to the primary visual cortex (V1). Similarly, the hierarchical structure proceeds from V1 into the secondary visual cortex (V2). Here, each level receives an input and estimates the object from its input. This estimate is calculated by a predictive estimator, learned from images the estimator is trained to, by a Kalman filter or generative system. It is argued in this model that this estimation is calculated *via* cortico-cortical connections in V1. Next, the model predicts the object at the next time step and conveys a predicted feature back to lower structures *via* descending cortical pathways. The usefulness of this hierarchical network model was established by its capacity to predict numerous types of neural and behavioral responses in the visual system. These include features such as: (1) distinguishing

a learned image from occluding objects (i.e., bottle partially occluding an image of a hand) and background noise added to the image; (2) predicting a sequence of images; (3) end stopping; and (4) other “extra-classical” receptive field effects (Rao and Ballard, 1999).

Establishing a neural implementation of predictive coding schemes has been challenging. At a minimum, one needs “prediction neurons” (or circuits) that provide a top-down signal and “prediction error neurons” (or circuits) that provide a bottom-up signal. In the context of the cerebral cortex, given the layer-specific directionality of cortical hierarchies (Rockland and Pandya, 1979; Felleman and Van Essen, 1991), prediction neurons would likely be found in the sources of descending connections: cells in layers 5 and 6. Since these cells project to layers 2 and 3 of areas lower in the sensory hierarchy, one would expect that supragranular layers would then contain prediction error neurons, as has been proposed previously (Bastos et al., 2012; Shipp, 2016). An important component of this basic circuit is the weighting of evidence from either bottom-up or top-down signals. For example, for highly reliable sensory signals, top-down predictions should carry less weight, while in situations of high sensory ambiguity (e.g., discerning a weak sound in noise), top-down signals should carry more weight. Most sensory systems do not have the luxury of repeatedly sampling the environment to determine the reliability of signals, but can estimate it based on saliency cues. It may be that neuromodulatory inputs (e.g., from cholinergic or monoaminergic fibers) can carry such a signal to dial up or down the reliance on top-down cues and thus adjust the “Kalman gain” of top-down modulation (Figure 3). Thus, sensory perception becomes a balance between reliance on top-down cues and bottom-up sensory saliency, as has recently been described human audition experiments (Huang and Elhilali, 2020). The over-reliance on top-down cues (possibly associated with disrupted neuromodulatory signals) may underlie pathophysiological states, such as the presence of delusions and hallucinations (Adams et al., 2014; Sterzer et al., 2018).

Physiological evidence for predictive coding at the single-neuron level has been observed in the visual cortex. Work in the late 1990s and early 2000s established that neurons in the early visual cortex of primates were sensitive to stimulus context and illusory signals (e.g., shape from shading or illusory contours in Kanizsa figures) and that these responses generally came after their response onsets (consistent with the time needed for feedback) and that the delayed responses were more characteristic of neurons from regions higher in the processing hierarchy (Lamme, 1995; Lee and Nguyen, 2001; Lee et al., 2002). Active silencing of descending connections from secondary visual areas can also eliminate surround suppressive effects, including end-stopping in V1 (Nassi et al., 2013), as proposed by Rao and Ballard (1999). More recent work has established similar patterns in the face-selective regions of the monkey temporal cortex (Schwiedrzik and Freiwald, 2017; Issa et al., 2018). In rodents, primary visual cortex neurons demonstrating responses to predictable stimuli, in advance of those stimuli, likely related to top-down signals from the cingulate cortex, have been identified



(Fiser et al., 2016). These data all suggest that neurons in both the early- and late-visual cortex receive inputs from higher regions in the visual hierarchy that confer inferential properties upon those neurons.

However, applying these or other physiological data to a predictive coding model faces several challenges. First, as outlined above, cortical connectivity patterns in the primate brain imply that prediction neurons should be found infragranularly, and prediction error neurons should be found supragranularly (as has been proposed; Bastos et al., 2012; Shipp, 2016). Accepting the notion that we could recognize a “prediction neuron” when we see it (Kogo and Trengove, 2015), it is not clear from the physiological literature that there are differences in prediction error sensitivity in the upper vs. lower layers of the auditory cortex. For example, Atencio and Schreiner observed marked differences between granular and not-granular layers in terms of their representation of sound across multiple dimensions in the cat, but no indication that prediction-type neurons resided in infragranular layers or that prediction error was represented supragranularly (Atencio et al., 2009). Another study that observed suppression of motor-related prediction signals found that prediction error signals were found to be represented in the deep layers (Rummell et al., 2016), which is the opposite of that described by current canonical models of predictive coding (Bastos et al., 2012). Second, the

general approach of subtracting away predictions implies that top-down projections should synapse on inhibitory neurons primarily—an idea for which there is little evidence—and that neural responses are smaller for predicted stimuli than for unpredicted stimuli. Regarding the former point, most work has revealed that descending intracortical projections form synapses on excitatory neurons and predominantly produce excitation (Johnson and Burkhalter, 1996; Shao and Burkhalter, 1996). Regarding the latter point, behavioral studies suggest that when ambiguous stimuli are congruent with expectations, behavioral performance is enhanced. Taken to its logical extent, the subtractive formulation of predictive coding implies that perfect predictions, which produce optimal behavior, *are associated with no neural responses*.

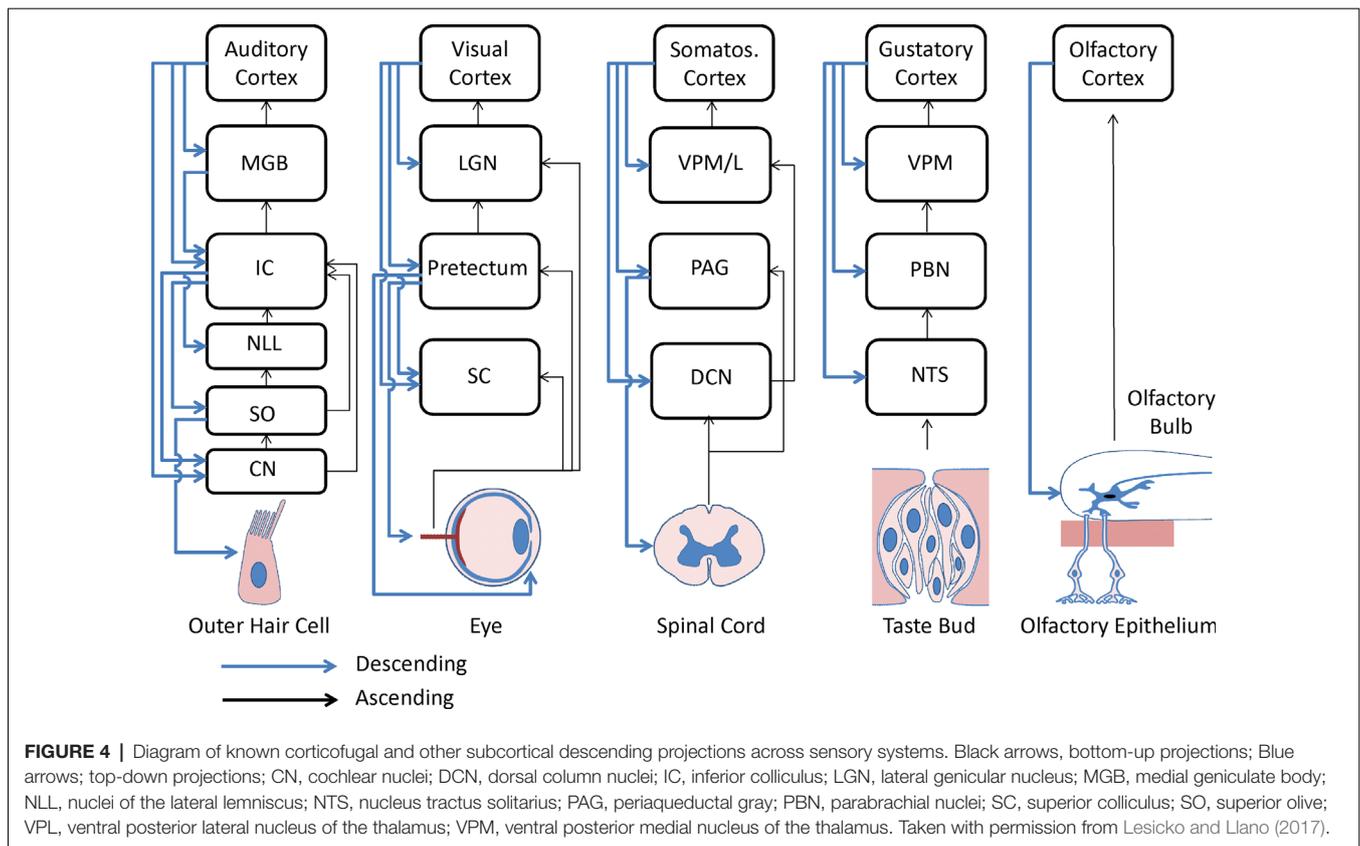
Most predictive coding schemes postulate that top-down predictions subtract from lower-level processors, leaving behind that which is not predicted—the prediction error. This scheme suggests that peripheral neurons are primarily responding to prediction errors—that which we do not predict. However, our behavior is just the opposite—we tend to ignore sensory data that do not fit into our predictions about the world. Thus, although predictive coding schemes that rely on the concept of prediction error can reproduce the responses of peripheral neurons, they do a poor job of explaining perception. We note that motor prediction may be a special case where subtraction is needed to remove the expected sensory consequences of actions (e.g., to suppress acoustic responses to vocalizations; Eliades and Wang, 2003), and here top-down motor-auditory circuits have been found to synapse on inhibitory interneurons (Nelson et al., 2013). More recent formulations have modified predictive coding algorithms to not include the subtraction operation for this reason (discussed in Spratling, 2017). Finally, predictive coding models have virtually ignored the massive sets of descending connections from the cortex that target subcortical regions, which have a very natural hierarchical organization. In the following sections, we explore the degree to which predictive coding models may be applied to the auditory corticofugal system.

EARLY VS. LATE TOP-DOWN MODULATION

As described above, most previous work on predictive coding in the auditory system has focused on the cerebral cortex. Corticocentric views of predictive coding have been driven by the fact that most of the relevant work on top-down modulation has been done in humans, where the techniques that are commonly used, EEG, MEG, and functional magnetic resonance imaging (fMRI), are most suited to measure activity in the cortex. Even though activity in subcortical structures may be seen in fMRI studies, they require appropriate hemodynamic response functions and often motion-correction procedures not needed for cortex, leading to the general absence of analysis of the subcortical activity in speech and language studies, as we have argued previously (Llano, 2013; Esmaeeli et al., 2019). However, there are massive projections to subcortical structures at all levels of the auditory system and these have been documented

for at least 100 years (Held, 1893). For example, in the visual system (the only system to our knowledge where such an analysis has been done) descending projections from the visual cortex outnumber ascending projections to the thalamus by at least 3-fold (Erişir et al., 1997). Beyond descending control to the thalamus, there are projections from the auditory cortex to the inferior colliculus (Fitzpatrick and Imig, 1978; Winer et al., 1998; Bajo and Moore, 2005; Bajo et al., 2007; Bajo and King, 2013; Torii et al., 2013; Stebbings et al., 2014), from the thalamus to the inferior colliculus (Kuwabara and Zook, 2000; Senatorov and Hu, 2002; Winer et al., 2002; Patel et al., 2017), from the inferior colliculus to the superior olive and cochlear nucleus (Conlee and Kane, 1982; Caicedo and Herbert, 1993; Saldaña, 1993; Vetter et al., 1993; Malmierca et al., 1996; Schofield, 2001; Groff and Liberman, 2003) and from the superior olive to the inner and outer hair cells in the cochlea (Liberman and Brown, 1986; Guinan, 2006). Thus, manipulations at the level of the auditory cortex, *via* these cascading descending projections, can, and have been shown to, substantially influence processing at the level of the cochlea (León et al., 2012). Indeed, early work established attentional effects at the level of single units in the cochlear nucleus in cats (Hernandez-Peon et al., 1956). Analogous projections from the sensory cortex to the sensory periphery have been identified in other sensory systems as well (see **Figure 4**), suggesting that early filtering in sensory systems may be a general principle for top-down modulation.

Other investigators have proposed potential advantages to the application of top-down modulation at the early (subcortical) processing stage, rather than later (cortical) processing stages (He, 2003a). For example, seminal work by Broadbent suggested an early filtering mechanism based on the apparent loss of information that was ignored during a dichotic listening task (Broadbent, 1958). Modifications to this theory to account for some retention of information filtered at an early stage were also proposed (Treisman, 1964). Most recently, a “new early filter model” was proposed by Marsh and Campbell (2016) which postulated that long-range corticofugal-corticopetal (ascending) loops may be responsible for early filtering of signals at the level of the brainstem (Marsh and Campbell, 2016) and that a tradeoff may exist between early and late filtering depending on task requirements. For example, very challenging attentional tasks or tasks that require very rapid processing of information may be better suited for an early filtering process (Giard et al., 2000). Also, tasks that require filtering based on features that are lost as information ascends the sensory hierarchy (e.g., fine temporal structure) may also be optimally filtered before those representations being lost (Marsh and Campbell, 2016). Importantly, however, top-down modulation in speech processing occurs at multiple levels of abstraction and at multiple time scales, some requiring higher-level filtering. For example, top-down information may come in the form of lexical cues (operating over ms) or prosodic cues (operating over ms to seconds) as well as other dimensions, such as using low-level cues such as voice familiarity vs. high-level pragmatic cues (Obleser, 2014). Thus, late (cortical) and early (subcortical) modulation may play complementary roles in top-down modulation during active listening.



METHODOLOGICAL ISSUES IN TOP-DOWN MODULATION IN THE SUBCORTICAL AUDITORY SYSTEM

Here we review methodological issues surrounding the study of descending projections from the auditory cortex to subcortical structures to effectuate top-down auditory control described above. It is worth noting that “descending projections” are not synonymous with top-down control. It is possible that lateral interactions within a brain structure (Srinivasan et al., 1982) can produce contextual modulation, as discussed in Rao and Ballard (1999) and Aitchison and Lengyel (2017). Here, we focus on corticofugal projections in keeping with the theme of this Special Issue on Cortical-Subcortical Loops in Sensory Processing.

Experimental paradigms for studying the corticofugal system have technical challenges that must be considered when analyzing the resulting data. Classical approaches include measuring response properties in a subcortical nucleus, then silencing the auditory cortex by cooling it or applying GABAergic agonists, and then re-measuring those properties. This paradigm is limited by: (1) incomplete recovery of cortical responses with certain GABAergic agents (Bauerle et al., 2011); (2) lack of specificity of which layer (layer 5 or layer 6 corticofugal neurons) is silenced; (3) lack of specificity about which frequency ranges across the tonotopic axis of the auditory cortex are silenced; and (4) lack of knowledge if the effects of silencing are on the brain structure being studied (e.g., thalamus or

inferior colliculus) or related to changes in the input to that structure from the cochlea, which is known to be impacted by cortical silencing (León et al., 2012). Regarding layer of origin, previous work has shown that both layers 5 and 6 project to the auditory thalamus and inferior colliculus (Games and Winer, 1988; Ojima, 1994; Künzle, 1995; Doucet et al., 2003; Bajo and Moore, 2005; Coomes et al., 2005; Llano and Sherman, 2008; Schofield, 2009; Slater et al., 2013, 2019), and that these projections have different physiological properties (Llano and Sherman, 2009; Slater et al., 2013) and likely different impacts on their target structures. Layer 5 cells have “driver”—type effects and layer 6 cells have “modulator”—type effects (for review see Lee and Sherman, 2010). Therefore bulk silencing is likely to homogenize the impacts of what could be quite different effects of these projections on their target structures. Likewise, work done using focal stimulation of the auditory cortex (reviewed in “Evidence That Auditory Corticofugal Systems Engage in Predictive Coding” section) suggests that corticofugal systems have markedly frequency-specific (in terms of the tonotopic axis) effects on their target structures, such that stimulation of neurons in certain frequency ranges can enhance, and others can suppress, subcortical responsiveness. Therefore, bulk silencing may produce a mixture of effects that are difficult to interpret. More modern approaches using viral-mediated delivery of optogenetic probes may solve some of these problems by permitting cell-type specific (Blackwell et al., 2020), layer-specific activation or silencing, and will permit activation

or silencing to occur at the level of terminals, diminishing the likelihood of indirect effects stemming from changes in cochlear function.

Activating corticofugal projections with electrical stimulation has also been used in many studies, but also has potential methodological pitfalls. Specific to the auditory thalamus, electrical stimulation may antidromically activate thalamocortical neurons, which may then activate other structures, such as the thalamic reticular nucleus, whose neurons project back to the dorsal thalamus, leading to indirect effects. Importantly, the specific protocol of electrical stimulation may make a large difference in the impact on subcortical neurons. Small changes in the relative timing of cortical vs. acoustic stimulation, relative amplitudes, pulse rates, etc. can change responses from excitatory to inhibitory, even with optogenetic stimulation (Guo et al., 2017; Vila et al., 2019). Also, many studies have used stimulation paradigms that are really perceptual learning paradigms. That is, by repeatedly stimulating the corticofugal fibers and observing a change in tuning in a target structure, one is no longer only studying on-line modulation of sensory responses based on prior knowledge, but instead is studying the impact of tetanic stimulation of corticofugal fibers on synaptic plasticity in the target structure. Finally, much of the early work done on corticofugal modulation has been done on anesthetized animals. We know from work in human subjects that top-down projections appear to be particularly vulnerable to anesthesia or other factors that alter consciousness (Boly et al., 2011; Raz et al., 2014; reviewed in Sikkens et al., 2019), and thus may not be adequately studied in an anesthetized animal.

EVIDENCE THAT AUDITORY CORTICOFUGAL SYSTEMS ENGAGE IN PREDICTIVE CODING

The auditory cortex sends massive projections to the auditory thalamus (and related thalamic reticular nucleus), the inferior colliculus, and the cochlear nucleus. The projections to the thalamus and inferior colliculus emanate from layers 5 and 6, while those to the cochlear nucleus appear to only emanate from layer 5. It is not yet known whether there is a single layer 5 system that projects to all subcortical nuclei, though evidence exists for the presence of individual layer 5 cells that branch to the auditory thalamus and inferior colliculus (Asokan et al., 2018). Early work suggests that the layer 5 projections to the inferior colliculus and cochlear nucleus are independent (Doucet et al., 2003), though it should be noted that the double-backlabel technique used in this study is prone to false negatives if the two tracers are not placed into physiologically-matched zones in each structure. The layer 6 projections to the auditory thalamus and inferior colliculus are likely at least partially independent since they are found in different sublayers of layer 6 (Llano and Sherman, 2008; Slater et al., 2013; Stebbings et al., 2014).

The auditory corticothalamic system is massive, develops early, before hearing onset (Torii et al., 2013), elicits responses in the majority of MGB neurons (Ryugo and Weinberger, 1976; Villa et al., 1991; He et al., 2002) that are strong enough to

induce immediate-early gene expression (Guo et al., 2007; Sun et al., 2007), produces both short (2 ms) and long (hundreds of milliseconds) latency responses (Serkov et al., 1976) and elicits both excitation (the dominant response in the lemniscal ventral subdivision) and inhibition (likely mediated *via* the thalamic reticular nucleus; Amato et al., 1969; He, 1997, 2003b; He et al., 2002; Xiong et al., 2004; Yu et al., 2004; Zhang et al., 2008). Activation of corticothalamic fibers can adjust tuning and sensitivity of auditory thalamic neurons (Guo et al., 2017) and appears to be critical for performance in perceptually-challenging tasks (Happel et al., 2014; Homma et al., 2017), as well as for directing plastic changes that occur in the thalamus (Zhang and Yan, 2008; Nelson et al., 2015). Importantly from the predictive coding perspective, corticothalamic projections appear to be organized topographically (Takayanagi and Ojima, 2006), such that cortical and thalamic areas that are matched for best frequency tend to produce corticothalamic excitation, while those that are unmatched tend to produce inhibition (He, 1997; He et al., 2002). Also, auditory thalamic neurons have been shown to be strongly sensitive to local stimulus predictability (Anderson et al., 2009; Antunes et al., 2010; Richardson et al., 2013; Cai et al., 2016), suggesting that they play a role in the coding of expectancy.

Several key experiments have been done to investigate the potential for corticothalamic fibers to contribute to predictive coding. One commonly-employed paradigm has been to apply repetitive stimulation of the auditory cortex to simulate a repeated acoustic motif and then to measure tuning properties to various parameters (sound frequency, combination-sensitivity, et cetera) before and after cortical stimulation. A consistent finding in the thalamus (and indeed in the inferior colliculus and cochlear nucleus, as described in the following paragraphs) is that stimulation of corticofugal fibers induces a shift of tuning of thalamic neurons towards the tuning of the particular region of the auditory cortex (so-called “egocentric selection”; Yan and Suga, 1996; Zhang et al., 1997; Zhang and Suga, 2000). From a Bayesian perspective, these data suggest that corticothalamic fibers contain “priors” such that the presence of highly prevalent stimuli (simulated by electrical cortical stimulation) makes it more likely that more peripheral responses in the thalamus, midbrain, or cochlear nucleus (i.e., posterior probabilities) are biased to respond more strongly to stimuli that are more likely to exist in the environment. The repeated stimulus presentation may be utilized to expand the cortical representation of Bayesian priors (Köver and Bao, 2010). As outlined in the “Methodological Issues in Top-Down Modulation in the Subcortical Auditory System” section, this paradigm falls short of establishing that corticothalamic fibers provide predictive coding signals because of the myriad problems with electrical cortical stimulation of the cortex, and because of the lack of establishment that acoustic stimuli use corticothalamic fibers to implement a predictive coding in the thalamus. Conversely, although it is well-established that training to alter the salience of an acoustic stimulus will shift neuronal tuning curves to be more responsive to that stimulus (Fritz et al., 2003), it remains to be established that the shift in tuning is caused by corticofugal projections.

An alternative approach has been to implement “surprise” paradigms, similar to MMN described in humans. The analogous finding at the single-unit level is known as stimulus-specific adaptation (SSA). In SSA, neurons diminish their responsiveness to repeated stimuli but retain their responsiveness to unexpected stimuli (Ulanovsky et al., 2003). Although it has been argued whether SSA is the neuronal-level instantiation of MMN (Farley et al., 2010; Carbajal and Malmierca, 2018), for our purposes, it is sufficient to state that SSA clearly reflects a key component of predictive coding: suppression of responses to predicted, presumably irrelevant stimuli. SSA has been established to exist in MGB neurons (Anderson et al., 2009; Antunes et al., 2010; Richardson et al., 2013; as well as neurons in the nonlemniscal inferior colliculus, below). Reversible silencing of corticothalamic fibers does not eliminate thalamic SSA, though it does alter other basic properties, suggesting that corticothalamic fibers play a strong role in modulating the thalamus, but may not confer SSA-sensitivity upon the thalamus (Antunes and Malmierca, 2011). We note that more aggressive nonreversible suppression diminishes thalamic SSA (Bäuerle et al., 2011), however, the significance of this finding is uncertain in the absence of reversibility of the cortical lesion.

We also note that the findings of SSA, and the paradigm employed by Suga and colleagues showing egocentric selection, are essentially orthogonal findings. That is, SSA represents the elimination of a predictable (presumably irrelevant) signal while the Suga paradigm represents the enhancement of a repeated, presumably behaviorally-important, signal. Evidence for both repetition suppression and repetition enhancement have been seen in the human subcortical auditory system (May and Tiitinen, 2010; Skoe and Kraus, 2010), though the latter is more in line with Bayesian notions of predictive coding. Thus, the data demonstrating egocentric shifts in thalamic receptive field properties suggest that corticothalamic projections may play an important role in providing a set of priors to thalamic neurons to bias their response properties, but may not be involved in repetition suppression manifesting as SSA.

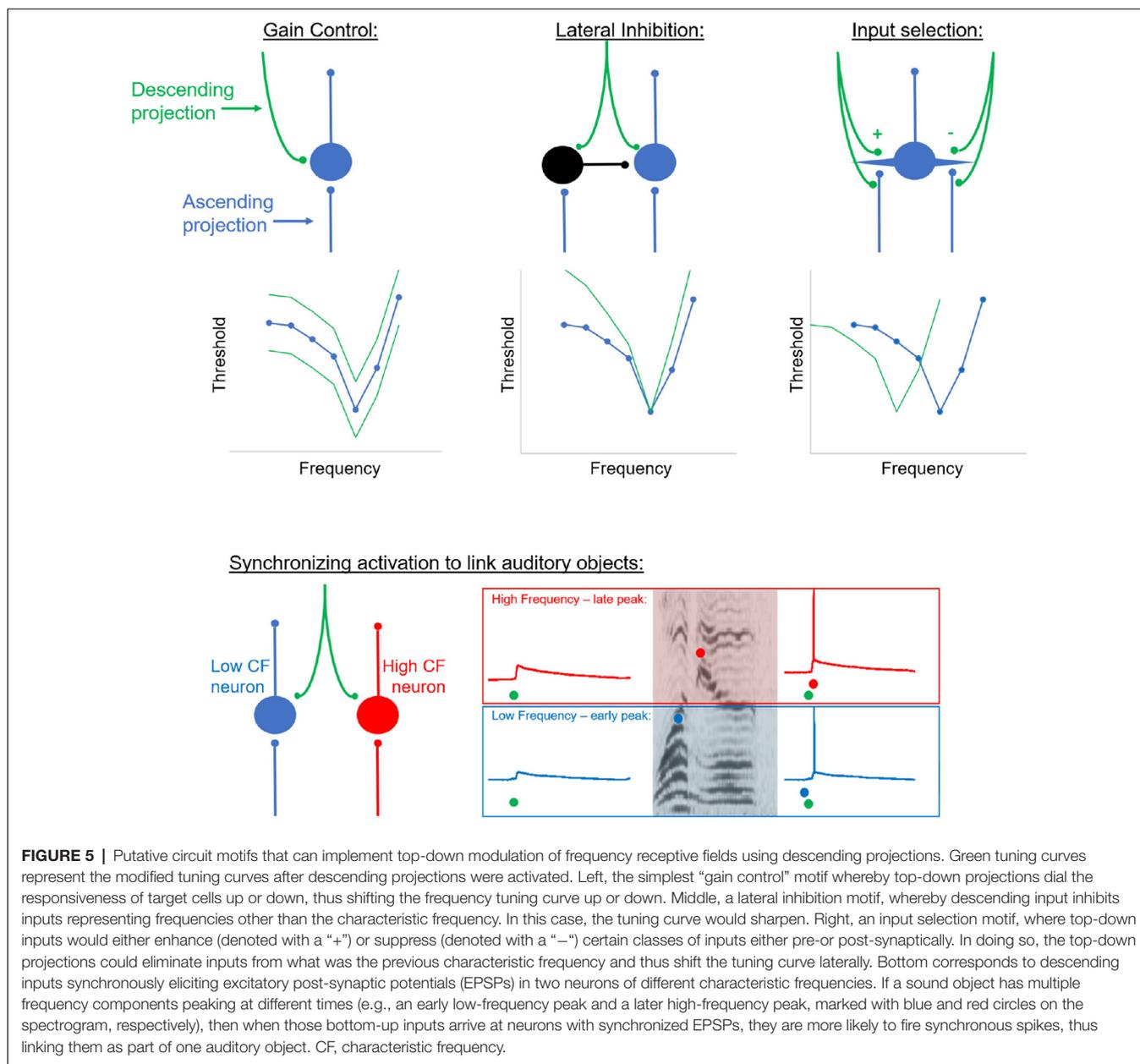
The corticocollicular system emanates primarily from layer 5 of the auditory cortex with a smaller component from layer 6 (Games and Winer, 1988; Künzle, 1995; Doucet et al., 2003; Bajo and Moore, 2005; Coomes et al., 2005; Schofield, 2009; Slater et al., 2013, 2019), and primarily targets the nonlemniscal portions of the inferior colliculus, grouped here as the lateral cortex and dorsal cortex (Saldaña et al., 1996; Winer et al., 1998). In the lateral cortex, the auditory projections interdigitate with somatosensory projections in a manner that is determined by neurochemical modules present in the lateral cortex (Lesicko et al., 2016). Electrical stimulation of the auditory cortex produces collicular responses with latencies as short as 1–2 ms (Mitani et al., 1983; Sun et al., 1989) and produces both excitation and inhibition (Mitani et al., 1983; Sun et al., 1989; Bledsoe et al., 2003; Markovitz et al., 2015). The projections are tonotopic (Lim and Anderson, 2007; Markovitz et al., 2013; Barnstedt et al., 2015) and the inhibition is presumably at least disynaptic because the corticocollicular system is thought to be excitatory (Feliciano and Potashner, 1995), and the suppression occurs in the later phases of the response (Popelář et al., 2015). Corticocollicular fibers

are responsible for protean functions at the level of the inferior colliculus, including facilitating adaptive changes in inferior colliculus neurons (Zhang et al., 2005; Wu and Yan, 2007; Bajo et al., 2010; Robinson et al., 2016; Asokan et al., 2018), sharpening of frequency tuning (Blackwell et al., 2020) and elicitation of escape responses (Xiong et al., 2015).

In terms of predictive coding, similar experiments to those done in the corticothalamic system have been done in the corticocollicular system but, in some cases, with a broader range of stimulus manipulations. For example, electrical stimulation of the auditory cortex causes egocentric shifts across multiple stimulus parameters, including frequency, duration, combination-sensitivity, sound location, and sound threshold (Yan and Suga, 1996, 1998; Jen et al., 1998; Ma and Suga, 2001; Yan and Ehret, 2001, 2002; Jen and Zhou, 2003; Yan et al., 2005; Zhou and Jen, 2005, 2007). These data suggest that the auditory cortex actively adjusts the tuning of collicular neurons to bias the response property across multiple computed stimulus dimensions and is not just inherited as part of the basic tonotopic layout of the two structures. Thus, a whole family of Bayesian priors (not unlike the family of hypotheses employed in particle filtering) can be used to modify the inferior colliculus. One challenge in understanding the corticocollicular findings is that most of the studies have involved recordings in the central nucleus of the inferior colliculus, which receives a small number of corticocollicular projections compared to the nonlemniscal regions. One potential resolution is that corticocollicular projections to the lateral cortex may have cascading inhibitory projections to the central nucleus after providing glutamatergic inputs to the lateral cortex, thus leading to primary inhibition in the central nucleus (Jen et al., 2001).

SSA has been observed in the dorsal and lateral cortices of the inferior colliculus (Malmierca et al., 2009; Duque et al., 2012), and it is thought that this is the earliest level that SSA occurs in the auditory system (Duque et al., 2018). Similar to the thalamus, reversible deactivation of the auditory cortex did not eliminate SSA in the inferior colliculus (Anderson and Malmierca, 2013). Thus, corticocollicular projections provide a strong predictive signal, possibly corraling inhibition from the lateral cortex en route to the central nucleus, to shift the tuning of collicular neurons towards those of previously heard stimuli. In contrast, suppression of repetitive irrelevant stimuli used in SSA appears to not involve these projections.

The auditory cortex also projects to the nuclei of the caudal auditory brainstem: cochlear nucleus, nucleus sagulum, and superior olivary nuclei (Feliciano and Potashner, 1995; Doucet et al., 2002; Meltzer and Ryugo, 2006), reviewed in Saldaña (2015). Compared to thalamic and collicular projections, comparatively little work has been done on these projections concerning predictive coding and all of it has been done in the cochlear nucleus. That said, all of the studies that have been done that measure tuning properties before and after focal cortical stimulation have revealed the same egocentric selection process described above for corticothalamic and corticocollicular neurons (Luo et al., 2008; Liu et al., 2010; Kong et al., 2014).



Notably, much of the early work on corticofugal modulation in animal models was done on echolocating bats (Yan and Suga, 1996, 1998; Zhang et al., 1997; Jen et al., 1998, 2001; Gao and Suga, 2000; Zhang and Suga, 2000; Ma and Suga, 2001). Although these mechanisms may be specific to echolocating bats due to their specialized behavioral requirements (Kössl et al., 2015), much of the key findings of the egocentric section have been seen in corticofugal projections non-echolocating species (Yan and Ehret, 2001, 2002; Yan et al., 2005; Luo et al., 2008; Liu et al., 2010; Kong et al., 2014). These data suggest that the basic principle of shifting tuning towards highly stimulated cortical representations is shared amongst both echolocating and non-echolocating species.

CIRCUIT-LEVEL MECHANISMS OF CORTICOFUGAL TOP-DOWN CONTROL

Virtually all work to date on corticofugal modulation in the auditory system has been done at the level of phenomenology without circuit-level analysis. Interestingly, corticothalamic, corticocollicular, and corticobulbar projections all appear to have similar effects on their targets—they produce egocentric modifications of receptive fields after repetitive stimulation. This similarity suggests a common neural substrate may exist across these projections. The layer 5 corticofugal system is common to these projections, and thus may be a potential candidate. Layer 5 corticofugal neurons have similar properties across regions of the cortex. They are large pyramidal cells with long and

tufted apical dendrites that burst intrinsically when depolarized (Connors et al., 1982; Kasper et al., 1994; Hefti and Smith, 2000; Llano and Sherman, 2009) and receive direct inputs from the thalamus (Constantinople and Bruno, 2013; Slater et al., 2019). In the corticothalamic system, these axons end in large terminals that synapse on proximal dendrites, producing “driver” type responses (Reichova and Sherman, 2004; Prasad et al., 2020). As described above, auditory corticothalamic terminals branch to the inferior colliculus (Asokan et al., 2018), but corticocollicular axons apparently do not branch to the cochlear nucleus (Doucet et al., 2003). In this respect, the layer 5 auditory corticothalamic system may diverge from other corticofugal systems where widespread subcortical branching is seen (Bourassa et al., 1995; Deschenes et al., 1996; Kita and Kita, 2012), reviewed in Usrey and Sherman (2019). Future work with sensitive tracers will clarify the extent to which a single auditory layer 5 “broadcast” neurons exist that send similar training signals to auditory thalamus, inferior colliculus and cochlear nucleus. Alternatively, given the homogenous nature of the changes seen across these three auditory nuclei, and the potential for auditory cortex stimulation to alter ascending information flow from the cochlea (León et al., 2012), these changes may be, in part, caused by alterations in shared ascending auditory information. We note that layer 6 projections to the thalamus are more numerous than layer 5 projections but tend to have smaller and more distal terminals (Lee and Sherman, 2010), and relay inhibition through the thalamic reticular nucleus (Lam and Sherman, 2010). Layer 6 corticocollicular projections also emanate from smaller neurons than layer 5 and have thinner neuronal projections and end in smaller terminals (Yudintsev et al., 2019). These data suggest that the layer 6 system may operate on a slower time scale, and is more likely to engage inhibitory interneurons, and thus may have a different set of functions than the layer 5 system that has yet to be identified.

The synaptic mechanisms by which auditory corticofugal projections modulate response properties are unknown, but several limitations based on previous extracellular recording studies exist. For example, to effectuate a change in tuning to sound frequency, a significantly more sophisticated operation than “gain control” must take place. To induce a neuron to respond to a frequency of sound to which it was not previously responsive over a matter of minutes, there must have existed a population of latent (i.e., inactive) inputs that are responsive to those frequencies. Conversely, a population of synapses encoding previously-responsive sounds would need to be silenced. Although inhibitory/disinhibitory mechanisms may create these types of shifts and do appear to play a role in the corticocollicular system, a small fraction of the corticocollicular system (4%) synapses on inhibitory interneurons (Nakamoto et al., 2013). An alternative mechanism could be to strengthen or weaken synapses without the use of inhibition. Repetitive, tetanic stimulation of a focal area of the auditory cortex has been well-established to alter receptive field properties of that area of the cortex (Ohl and Scheich, 1997, 2005; Weinberger, 2004). Repetitive acoustic stimulation may also decrease the representation of that sound in the auditory cortex, depending on the behavioral salience of that sound (Condon and Weinberger,

1991). It is therefore possible that descending connections could strengthen synapses post-synaptically, though in the absence of an appropriately timed ascending signal would appear to be a non-Hebbian mechanism to induce a plastic change. Descending projections could also target presynaptic terminals to either activate them or diminish their strength, as suggested by early work in the visual system (Iwama et al., 1965), see **Figure 5**. However, at least in the auditory corticocollicular system, little evidence for presynaptic terminals exists in the corticocollicular or auditory corticothalamic system (Bartlett et al., 2000; Nakamoto et al., 2013). Beyond impacts at the level of individual cells, corticofugal projections may influence a population of cells to alter their likelihood of firing synchronously, as proposed previously (Gilbert and Li, 2013). Such a mechanism would be ideally suited to either integrate disparate pieces of information (as needed for contour integration, or phonemic restoration) or to segregate information (as needed during speech segmentation or stream segregation). For example, neural responses to a sound object with complex spectrotemporal properties with low and high-frequency peaks at different times may be linked into a singular perceptual object if descending projections synchronized subthreshold responses across an array of sensory neurons (**Figure 5**, bottom). Thus, unsynchronized responses from neurons with different characteristic frequencies at low levels of the hierarchy could be tagged as being derived from the same acoustic object by eliciting synchronized responses at higher levels of the hierarchy. Very little work of this type has been done, though it should be noted that inhibition of the corticothalamic system leads to greater synchrony of firing between thalamic neurons, suggesting that the corticothalamic system has the potential to enhance segregation between input streams (Villa et al., 1999). Similar findings were reported in the corticocollicular system by Nakamoto et al. (2010). Thus, multiple non-mutually exclusive synaptic motifs may help to explain the impact(s) of the corticofugal systems, and none have been systematically explored to date.

CONCLUSIONS AND FUTURE CHALLENGES

Top-down modulation is observed at the level of behavior and the level of the single neurons, and there is still much work left to be done to understand how these two levels of top-down modulation are linked. In our view, the weight of the evidence suggests that at least one role of descending projections is to modify receptive field properties to bias them towards frequently-occurring or highly salient stimuli. However, consistent with the anatomical and physiological heterogeneity of these systems, additional roles are possible. Complicating matters is the finding that these systems are often intermingled and individual projections may have more than one role. A challenge, then, in the field is how to design an experimental paradigm to identify the circuit motifs that produce top-down modulation and how they alter perceptual responses. The first step is to decide precisely what is being studied. The term “predictive coding” is broad enough to encompass many different types of processing. For example, the term is used to

describe both the “explaining away” of expected and ignored stimuli as well as the enhancement of expected but obscured stimuli. As described above, the computations underlying these two processes are not the same and do not appear to be handled by the same circuits. The other challenging experimental question is deciding which level of top-down modulation is to be studied.

There are many “descending systems” in the auditory system and many types of tasks that require top-down modulation. Descending projections extend from the frontal cortex to the auditory cortex and on to the cochlea, with a stop at every auditory subcortical structure along the way. Presumably, certain descending projections should be important for high-level modulation (e.g., using discourse cues to understand an ambiguous word) vs. low-level modulation (e.g., having a loud sound diminish the sensitivity of the cochlea to subsequent sounds). An additional dimension is task difficulty. That is, difficult tasks may require multiple descending projections to be involved, thereby altering the stimulus representation as soon as it enters the brain, and others may be less challenging, allowing later filtering, thus permitting several stimulus representations to “coexist” in the brain before one being selected. Therefore, engaging in a systematic process to identify which pathway is engaged during which task would be a starting point for future investigators.

Also, to facilitate comparisons across studies, it will be important for future experiments to specify the type and level of predictive coding being studied. Also, although electrical stimulation paradigms have provided insights about predictive coding by demonstrating that repetitive activation of a particular region of cortex can change the filtering properties of more peripheral sensory neurons (reviewed above), these changes typically have been found after long-term (minutes) tetanic stimulation of the corticofugal projections, which is a crude approximation to altering the statistical likelihood of a particular sound appearing in the environment. A more convincing demonstration would be to show that the tuning of a particular neuron changes dynamically, and under particular behavioral contexts (similar to that seen in Caras and Sanes, 2017) when the likelihood that a particular stimulus occurs changes. Besides, one would also anticipate that prediction neurons would have their strongest impact when peripheral signals are weak (i.e., the Kalman gain would be highest under these circumstances). Consistent with this idea, previous work has shown that top-down modulation tends to be strongest in broadly tuned neurons [presumably neurons with ambiguous frequency representations (Vila et al., 2019) or when acoustic stimulus amplitude is weak (Jen et al., 1998)]. It may be the

case that neurons that are broadly tuned to isolated sounds may be more sharply tuned in other contexts. Also, future work should emphasize paradigms that alter stimulus expectancy without altering stimulus probability in awake animals [as used in Cai et al. (2016)], thus removing the bottom-up cue of stimulus probability. Finally, one experimentally pragmatic benefit of studying corticofugal systems is the physical separation between the descending system and the physical structure under study, allowing the examination of responses in putative “prediction axons” (presumably corticofugal) compared to bottom-up signals. This type of approach has been used in two-photon imaging of the visual system, where presumed prediction neurons in the anterior cingulate were labeled and their response properties appeared to carry prediction signals (Fiser et al., 2016). The use of this set of approaches would get us closer to understanding the unusual connectivity patterns described by Lorente de Nó almost 100 years ago (Lorente De Nó, 1933):

“The conception of the reflex arc as a unidirectional chain of neurons has neither anatomic nor functional basis. Histologic studies. . . show the universality of the existence of plural parallel connections and of recurrent, reciprocal connections.”

Thus, a deliberate approach using techniques to interrogate populations of neurons in awake animals will permit the understanding of the logic of highly recurrent systems whose roles have remained obscure for nearly a century.

AUTHOR CONTRIBUTIONS

AA and DL both wrote the manuscript together. All authors contributed to the article and approved the submitted version.

FUNDING

DL was supported by National Institutes of Health DC013073 and AG059103.

ACKNOWLEDGMENTS

We thank Drs. Donald Caspary, Manuel Malmierca, Tom Anastasio, Silvio Macias, and Brett Schofield for productive discussions about the contents of this review and Dr. Nathiya Vaithiyalingam Chandra Sekaran for translating assistance for the Held (1893) reference. We also thank Drs. Lionel Collet, Paul Délano, Xavier Perrot, and David Smith for permission and/or supplying high-resolution versions of their figures.

REFERENCES

- Adams, R. A., Brown, H. R., and Friston, K. J. (2014). Bayesian inference, predictive coding and delusions. *Avant* 5, 51–88. doi: 10.26913/50302014.0112.0004
- Aitchison, L., and Lengyel, M. (2017). With or without you: predictive coding and bayesian inference in the brain. *Curr. Opin. Neurobiol.* 46, 219–227. doi: 10.1016/j.conb.2017.08.010
- Amato, G., La, V. G., and Enia, F. (1969). The control exerted by the auditory cortex on the activity of the medial geniculate body and inferior colliculus. *Arch. Sci. Biol.* 53, 291–313.
- Anderson, L. A., Christianson, G. B., and Linden, J. F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *J. Neurosci.* 29, 7359–7363. doi: 10.1523/JNEUROSCI.0793-09.2009
- Anderson, L. A., and Malmierca, M. (2013). The effect of auditory cortex deactivation on stimulus-specific adaptation in the inferior colliculus of the rat. *Eur. J. Neurosci.* 37, 52–62. doi: 10.1111/ejn.12018

- Andersson, L., Sandberg, P., Olofsson, J. K., and Nordin, S. (2018). Effects of task demands on olfactory, auditory and visual event-related potentials suggest similar top-down modulation across senses. *Chem. Senses* 43, 129–134. doi: 10.1093/chemse/bjx082
- Antunes, F. M., and Malmierca, M. S. (2011). Effect of auditory cortex deactivation on stimulus-specific adaptation in the medial geniculate body. *J. Neurosci.* 31, 17306–17316. doi: 10.1523/JNEUROSCI.1915-11.2011
- Antunes, F. M., Nelken, I., Covey, E., and Malmierca, M. S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS One* 5:e14071. doi: 10.1371/journal.pone.0014071
- Asokan, M. M., Williamson, R. S., Hancock, K. E., and Polley, D. B. (2018). Sensory overamplification in layer 5 auditory corticofugal projection neurons following cochlear nerve synaptic damage. *Nat. Commun.* 9:2468. doi: 10.1038/s41467-018-04852-y
- Atencio, C. A., Sharpee, T. O., and Schreiner, C. E. (2009). Hierarchical computation in the canonical auditory cortical circuit. *Proc. Natl. Acad. Sci. U S A* 106, 21894–21899. doi: 10.1073/pnas.0908383106
- Bajo, V. M., and King, A. J. (2013). Cortical modulation of auditory processing in the midbrain. *Front. Neural Circuits* 6:114. doi: 10.3389/fncir.2012.00114
- Bajo, V. M., and Moore, D. R. (2005). Descending projections from the auditory cortex to the inferior colliculus in the gerbil, *Meriones unguiculatus*. *J. Comp. Neurol.* 486, 101–116. doi: 10.1002/cne.20542
- Bajo, V. M., Nodal, F. R., Bizley, J. K., Moore, D. R., and King, A. J. (2007). The ferret auditory cortex: descending projections to the inferior colliculus. *Cereb. Cortex* 17, 475–491. doi: 10.1093/cercor/bhj164
- Bajo, V. M., Nodal, F. R., Moore, D. R., and King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13, 253–260. doi: 10.1038/nn.2466
- Barnstedt, O., Keating, P., Weissenberger, Y., King, A. J., and Dahmen, J. C. (2015). Functional microarchitecture of the mouse dorsal inferior colliculus revealed through *in vivo* two-photon calcium imaging. *J. Neurosci.* 35, 10927–10939. doi: 10.1523/JNEUROSCI.0103-15.2015
- Bartlett, E. L., Stark, J. M., Gullery, R. W., and Smith, P. H. (2000). Comparison of the fine structure of cortical and collicular terminals in the rat medial geniculate body. *Neuroscience* 100, 811–828. doi: 10.1016/s0306-4522(00)00340-7
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038
- Battaglia, P. W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1391–1397. doi: 10.1364/josaa.20.001391
- Bäuerle, P., Von Der Behrens, W., Kössl, M., and Gaese, B. H. (2011). Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *J. Neurosci.* 31, 9708–9722. doi: 10.1523/JNEUROSCI.5814-10.2011
- Blackwell, J. M., Lesicko, A. M. H., Rao, W., De Biasi, M., and Geffen, M. N. (2020). Auditory cortex shapes sound responses in the inferior colliculus. *eLife* 9:e51890. doi: 10.7554/eLife.51890
- Bledsoe, S., Shore, S. E., and Guitton, M. (2003). Spatial representation of corticofugal input in the inferior colliculus: a multicontact silicon probe approach. *Exp. Brain Res.* 153, 530–542. doi: 10.1007/s00221-003-1671-6
- Boly, M., Garrido, M. I., Gosseries, O., Bruno, M.-A., Boveroux, P., Schnakers, C., et al. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science* 332, 858–862. doi: 10.1126/science.1202043
- Bourassa, J., Pinault, D., and Deschênes, M. (1995). Corticothalamic projections from the cortical barrel field to the somatosensory thalamus in rats: a single-fibre study using biocytin as an anterograde tracer. *Eur. J. Neurosci.* 7, 19–30. doi: 10.1111/j.1460-9568.1995.tb01016.x
- Braga, R. M., Wilson, L. R., Sharp, D. J., Wise, R. J., and Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. *NeuroImage* 74, 77–86. doi: 10.1016/j.neuroimage.2013.02.023
- Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT press.
- Broadbent, D. (1958). *Perception and Communication*. Elmsford, NY: Pergamon Press.
- Cai, R., Richardson, B. D., and Caspary, D. M. (2016). Responses to predictable versus random temporally complex stimuli from single units in auditory thalamus: impact of aging and anesthesia. *J. Neurosci.* 36, 10696–10706. doi: 10.1523/JNEUROSCI.1454-16.2016
- Caicedo, A., and Herbert, H. (1993). Topography of descending projections from the inferior colliculus to auditory brainstem nuclei in the rat. *J. Comp. Neurol.* 328, 377–392. doi: 10.1002/cne.903280305
- Caras, M. L., and Sanes, D. H. (2017). Top-down modulation of sensory cortex gates perceptual learning. *Proc. Natl. Acad. Sci. U S A* 114, 9972–9977. doi: 10.1073/pnas.1712305114
- Carbajal, G. V., and Malmierca, M. S. (2018). The neuronal basis of predictive coding along the auditory pathway: from the subcortical roots to cortical deviance detection. *Trends Hear.* 22:2331216518784822. doi: 10.1177/2331216518784822
- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., and Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron* 64, 311–319. doi: 10.1016/j.neuron.2009.10.006
- Chen, Z. (2003). Bayesian filtering: from kalman filters to particle filters and beyond. *Statistics* 182, 1–69. doi: 10.1080/02331880309257
- Chennu, S., Noreika, V., Gueorguiev, D., Blenkman, A., Kochen, S., Ibáñez, A., et al. (2013). Expectation and attention in hierarchical auditory prediction. *J. Neurosci.* 33, 11194–11205. doi: 10.1523/JNEUROSCI.0114-13.2013
- Chennu, S., Noreika, V., Gueorguiev, D., Shtyrov, Y., Bekinschtein, T. A., and Henson, R. (2016). Silent expectations: dynamic causal modeling of cortical prediction and attention to sounds that weren't. *J. Neurosci.* 36, 8305–8316. doi: 10.1523/JNEUROSCI.1125-16.2016
- Collet, L., Bouchet, P., and Pernier, J. (1994). Auditory selective attention in the human cochlea. *Brain Res.* 633, 353–356. doi: 10.1016/0006-8993(94)91561-x
- Condon, C. D., and Weinberger, N. M. (1991). Habituation produces frequency-specific plasticity of receptive fields in the auditory cortex. *Behav. Neurosci.* 105, 416–430. doi: 10.1037/0735-7044.105.3.416
- Conlee, J., and Kane, E. (1982). Descending projections from the inferior colliculus to the dorsal cochlear nucleus in the cat: an autoradiographic study. *Neuroscience* 7, 161–178. doi: 10.1016/0306-4522(82)90158-0
- Connors, B. W., Gutnick, M. J., and Prince, D. A. (1982). Electrophysiological properties of neocortical neurons *in vitro*. *J. Neurophysiol.* 48, 1302–1320. doi: 10.1152/jn.1982.48.6.1302
- Constantinople, C. M., and Bruno, R. M. (2013). Deep cortical layers are activated directly by thalamus. *Science* 340, 1591–1594. doi: 10.1126/science.1236425
- Coomes, D. L., Schofield, R. M., and Schofield, B. R. (2005). Unilateral and bilateral projections from cortical cells to the inferior colliculus in guinea pigs. *Brain Res.* 1042, 62–72. doi: 10.1016/j.brainres.2005.02.015
- Cope, T. E., Sohoglu, E., Sedley, W., Patterson, K., Jones, P., Wiggins, J., et al. (2017). Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nat. Commun.* 8:2154. doi: 10.1038/s41467-017-01958-7
- Cunillera, T., Càmarà, E., Laine, M., and Rodríguez-Fornells, A. (2010). Words as anchors: known words facilitate statistical learning. *Exp. Psychol.* 57, 134–141. doi: 10.1027/1618-3169/a000017
- Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147. doi: 10.1016/j.heares.2007.01.014
- De Boer, J., and Thornton, A. R. D. (2008). Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J. Neurosci.* 28, 4929–4937. doi: 10.1523/JNEUROSCI.0902-08.2008
- Deschênes, M., Bourassa, J., Doan, V. D., and Parent, A. (1996). A single-cell study of the axonal projections arising from the posterior intralaminar thalamic nuclei in the rat. *Eur. J. Neurosci.* 8, 329–343. doi: 10.1111/j.1460-9568.1996.tb01217.x
- Doucet, J., Molavi, D., and Ryugo, D. (2003). The source of corticocollicular and corticobulbar projections in area Te1 of the rat. *Exp. Brain Res.* 153, 461–466. doi: 10.1007/s00221-003-1604-4

- Doucet, J. R., Rose, L., and Ryugo, D. K. (2002). The cellular origin of corticofugal projections to the superior olivary complex in the rat. *Brain Res.* 925, 28–41. doi: 10.1016/S0006-8993(01)03248-6
- Dragicevic, C. D., Aedo, C., León, A., Bowen, M., Jara, N., Terreros, G., et al. (2015). The olivocochlear reflex strength and cochlear sensitivity are independently modulated by auditory cortex microstimulation. *J. Assoc. Res. Otolaryngol.* 16, 223–240. doi: 10.1007/s10162-015-0509-9
- Duque, D., Pais, R., and Malmierca, M. S. (2018). Stimulus-specific adaptation in the anesthetized mouse revealed by brainstem auditory evoked potentials. *Hear. Res.* 370, 294–301. doi: 10.1016/j.heares.2018.08.011
- Duque, D., Pérez-González, D., Ayala, Y. A., Palmer, A. R., and Malmierca, M. S. (2012). Topographic distribution, frequency and intensity dependence of stimulus-specific adaptation in the inferior colliculus of the rat. *J. Neurosci.* 32, 17762–17774. doi: 10.1523/JNEUROSCI.3190-12.2012
- Eliades, S. J., and Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* 89, 2194–2207. doi: 10.1152/jn.00627.2002
- Erişir, A., Van Horn, S. C., and Sherman, S. M. (1997). Relative numbers of cortical and brainstem inputs to the lateral geniculate nucleus. *Proc. Natl. Acad. Sci. U S A* 94, 1517–1520. doi: 10.1073/pnas.94.4.1517
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Esmaeli, S., Murphy, K., Swords, G. M., Ibrahim, B. A., Brown, J. W., and Llano, D. A. (2019). Visual hallucinations, thalamocortical physiology and Lewy body disease: a review. *Neurosci. Biobehav. Rev.* 103, 337–351. doi: 10.1016/j.neubiorev.2019.06.006
- Farley, B. J., Quirk, M. C., Doherty, J. J., and Christian, E. P. (2010). Stimulus-specific adaptation in auditory cortex is an NMDA-independent process distinct from the sensory novelty encoded by the mismatch negativity. *J. Neurosci.* 30, 16475–16484. doi: 10.1523/JNEUROSCI.2793-10.2010
- Feliciano, M., and Potashner, S. J. (1995). Evidence for a glutamatergic pathway from the guinea pig auditory cortex to the inferior colliculus. *J. Neurochem.* 65, 1348–1357. doi: 10.1046/j.1471-4159.1995.65031348.x
- Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1
- Fink, M., Churan, J., and Wittmann, M. (2006). Temporal processing and context dependency of phoneme discrimination in patients with aphasia. *Brain Lang.* 98, 1–11. doi: 10.1016/j.bandl.2005.12.005
- Fiser, A., Mahringer, D., Oyibo, H. K., Petersen, A. V., Leinweber, M., and Keller, G. B. (2016). Experience-dependent spatial expectations in mouse visual cortex. *Nat. Neurosci.* 19, 1658–1664. doi: 10.1038/nn.4385
- Fitzpatrick, K. A., and Imig, T. J. (1978). Projections of auditory cortex upon the thalamus and midbrain in the owl monkey. *J. Comp. Neurol.* 177, 537–555. doi: 10.1002/cne.901770402
- Fodor, J. A., and Bever, T. G. (1965). The psychological reality of linguistic segments. *J. Verbal Learning Verbal Behav.* 4, 414–420. doi: 10.1016/S0022-5371(65)80081-0
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J. Acoust. Soc. Am.* 115, 2246–2256. doi: 10.1121/1.1689343
- Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1211–1221. doi: 10.1098/rstb.2008.0300
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223. doi: 10.1038/nn1141
- Games, K. D., and Winer, J. A. (1988). Layer V in rat auditory cortex: projections to the inferior colliculus and contralateral cortex. *Hear. Res.* 34, 1–25. doi: 10.1016/0378-5955(88)90047-0
- Ganong, W. F. III (1980). Phonetic categorization in auditory word perception. *J. Exp. Psychol. Hum. Percept. Perform.* 6, 110–125. doi: 10.1037/0096-1523.6.1.110
- Gao, E., and Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc. Natl. Acad. Sci. U S A* 97, 8081–8086. doi: 10.1073/pnas.97.14.8081
- García-Rosales, F., López-Jury, L., González-Palomares, E., Cabral-Calderín, Y., and Hechavarría, J. C. (2020). Fronto-temporal coupling dynamics during spontaneous activity and auditory processing in the bat *Carollia perspicillata*. *Front. Syst. Neurosci.* 14:14. doi: 10.1016/j.biologicals.2020.01.008
- Getz, L. M., and Toscano, J. C. (2019). Electrophysiological evidence for top-down lexical influences on early speech perception. *Psychol. Sci.* 30, 830–841. doi: 10.1177/0956797619841813
- Giard, M.-H., Fort, A., Mouchetant-Rostaing, Y., and Pernier, J. (2000). Neurophysiological mechanisms of auditory selective attention in humans. *Front. Biosci.* 5, D84–D94. doi: 10.2741/giard
- Gilbert, C. D., and Li, W. (2013). Top-down influences on visual processing. *Nat. Rev. Neurosci.* 14, 350–363. doi: 10.1038/nrn3476
- Grindrod, C. M., and Baum, S. R. (2002). Sentence context effects and the timecourse of lexical ambiguity resolution in nonfluent aphasia. *Brain Cogn.* 48, 381–385.
- Groff, J. A., and Liberman, M. C. (2003). Modulation of cochlear afferent response by the lateral olivocochlear system: activation via electrical stimulation of the inferior colliculus. *J. Neurophysiol.* 90, 3178–3200. doi: 10.1152/jn.00537.2003
- Guinan, J. J. Jr. (2006). Olivocochlear efferents: anatomy, physiology, function and the measurement of efferent effects in humans. *Ear Hear.* 27, 589–607. doi: 10.1097/01.aud.0000240507.83072.e7
- Guo, W., Clause, A. R., Barth-Maron, A., and Polley, D. B. (2017). A corticothalamic circuit for dynamic switching between feature detection and discrimination. *Neuron* 95, 180.e5–194.e5. doi: 10.1016/j.neuron.2017.05.019
- Guo, Y. P., Sun, X., Li, C., Wang, N. Q., Chan, Y.-S., and He, J. (2007). Corticothalamic synchronization leads to c-fos expression in the auditory thalamus. *Proc. Natl. Acad. Sci. U S A* 104, 11802–11807. doi: 10.1073/pnas.0701302104
- Haegens, S., Händel, B. F., and Jensen, O. (2011). Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *J. Neurosci.* 31, 5197–5204. doi: 10.1523/JNEUROSCI.5199-11.2011
- Hannemann, R., Obleser, J., and Eulitz, C. (2007). Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Res.* 1153, 134–143. doi: 10.1016/j.brainres.2007.03.069
- Happel, M. F., Deliano, M., Handschuh, J., and Ohl, F. W. (2014). Dopamine-modulated recurrent corticoefferent feedback in primary sensory cortex promotes detection of behaviorally relevant stimuli. *J. Neurosci.* 34, 1234–1247. doi: 10.1523/JNEUROSCI.1990-13.2014
- He, J. (1997). Modulatory effects of regional cortical activation on the onset responses of the cat medial geniculate neurons. *J. Neurophysiol.* 77, 896–908. doi: 10.1152/jn.1997.77.2.896
- He, J. (2003a). Corticofugal modulation of the auditory thalamus. *Exp. Brain Res.* 153, 579–590. doi: 10.1007/s00221-003-1680-5
- He, J. (2003b). Corticofugal modulation on both on and off responses in the nonlemniscal auditory thalamus of the guinea pig. *J. Neurophysiol.* 89, 367–381. doi: 10.1152/jn.00593.2002
- He, J., Yu, Y.-Q., Xiong, Y., Hashikawa, T., and Chan, Y.-S. (2002). Modulatory effect of cortical activation on the lemniscal auditory thalamus of the guinea pig. *J. Neurophysiol.* 88, 1040–1050. doi: 10.1152/jn.2002.88.2.1040
- Hefli, B. J., and Smith, P. H. (2000). Anatomy, physiology and synaptic responses of rat layer V auditory cortical cells and effects of intracellular GABA_A blockade. *J. Neurophysiol.* 83, 2626–2638. doi: 10.1152/jn.2000.83.5.2626
- Held, H. (1893). Die centrale gehörleitung. *Arch. Anat. Physiol. Anat. Abt.* 201–248
- Hernandez-Peon, R., Scherrer, H., and Jouve, M. (1956). Modification of electric activity in cochlear nucleus during attention in unanesthetized cats. *Science* 123, 331–332. doi: 10.1126/science.123.3191.331
- Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science* 182, 177–180. doi: 10.1126/science.182.4108.177
- Hofmann-Shen, C., Vogel, B. O., Kaffes, M., Rudolph, A., Brown, E. C., Tas, C., et al. (2020). Mapping adaptation, deviance detection and prediction error in auditory processing. *NeuroImage* 207:116432. doi: 10.1016/j.neuroimage.2019.116432
- Homma, N. Y., Happel, M. F., Nodal, F. R., Ohl, F. W., King, A. J., and Bajo, V. M. (2017). A role for auditory corticothalamic feedback in the perception of

- complex sounds. *J. Neurosci.* 37, 6149–6161. doi: 10.1523/JNEUROSCI.0397-17.2017
- Huang, N., and Elhilali, M. (2020). Push-pull competition between bottom-up and top-down auditory attention to natural soundscapes. *eLife* 9:e52984. doi: 10.7554/eLife.52984
- Issa, E. B., Cadieu, C. F., and Dicarlo, J. J. (2018). Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *eLife* 7:e42870. doi: 10.7554/eLife.42870
- Iwama, K., Sakajura, H., and Kasamatsu, T. (1965). Presynaptic inhibition in the lateral geniculate body induced by stimulation of the cerebral cortex. *Jap. J. Physiol.* 15, 310–322. doi: 10.2170/jjphysiol.15.310
- Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vis. Res.* 39, 3621–3629. doi: 10.1016/s0042-6989(99)00088-7
- Jäger, K., and Kössl, M. (2016). Corticofugal modulation of DPOAEs in gerbils. *Hear. Res.* 332, 61–72. doi: 10.3390/j3030024
- Jen, P. H.-S., Chen, Q. C., and Sun, X. D. (1998). Corticofugal regulation of auditory sensitivity in the bat inferior colliculus. *J. Comp. Physiol. A* 183, 683–697. doi: 10.1007/s003590050291
- Jen, P. H.-S., Sun, X. D., and Chen, Q. C. (2001). An electrophysiological study of neural pathways for corticofugally inhibited neurons in the central nucleus of the inferior colliculus of the big brown bat, *Eptesicus fuscus*. *Exp. Brain Res.* 137, 292–302. doi: 10.1007/s002210000637
- Jen, P. H.-S., and Zhou, X. (2003). Corticofugal modulation of amplitude domain processing in the midbrain of the big brown bat, *Eptesicus fuscus*. *Hear. Res.* 184, 91–106. doi: 10.1016/s0378-5955(03)00237-5
- Johnson, R. R., and Burkhalter, A. (1996). Microcircuitry of forward and feedback connections within rat visual cortex. *J. Comp. Neurol.* 368, 383–398. doi: 10.1002/(SICI)1096-9861(19960506)368:3<383::AID-CNE5>3.0.CO;2-1
- Jones, J. A., and Freyman, R. L. (2012). Effect of priming on energetic and informational masking in a same-different task. *Ear Hear.* 33, 124–133. doi: 10.1097/AUD.0b013e3182b25bee
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82, 35–45. doi: 10.1115/1.3662552
- Kasper, E. M., Larkman, A. U., Lübke, J., and Blakemore, C. (1994). Pyramidal neurons in layer 5 of the rat visual cortex. I. Correlation among cell morphology, intrinsic electrophysiological properties and axon targets. *J. Comp. Neurol.* 339, 459–474. doi: 10.1002/cne.903390402
- Kita, T., and Kita, H. (2012). The subthalamic nucleus is one of multiple innervation sites for long-range corticofugal axons: a single-axon tracing study in the rat. *J. Neurosci.* 32, 5990–5999. doi: 10.1523/JNEUROSCI.5717-11.2012
- Kobayashi, M., Takeda, M., Hattori, N., Fukunaga, M., Sasabe, T., Inoue, N., et al. (2004). Functional imaging of gustatory perception and imagery: “top-down” processing of gustatory signals. *NeuroImage* 23, 1271–1282. doi: 10.1016/j.neuroimage.2004.08.002
- Kogo, N., and Trengove, C. (2015). Is predictive coding theory articulated enough to be testable? *Front. Comput. Neurosci.* 9:111. doi: 10.3389/fncom.2015.00111
- Kong, L., Xiong, C., Li, L., and Yan, J. (2014). Frequency-specific corticofugal modulation of the dorsal cochlear nucleus in mice. *Front. Syst. Neurosci.* 8:125. doi: 10.3389/fnsys.2014.00125
- Kössl, M., Hechavarría, J., Voss, C., Schaefer, M., and Vater, M. (2015). Bat auditory cortex-model for general mammalian auditory computation or special design solution for active time perception? *Eur. J. Neurosci.* 41, 518–532. doi: 10.1111/ejn.12801
- Köver, H., and Bao, S. (2010). Cortical plasticity as a mechanism for storing Bayesian priors in sensory perception. *PLoS One* 5:e10497. doi: 10.1371/journal.pone.0010497
- Künzle, H. (1995). Regional and laminar distribution of cortical neurons projecting to either superior or inferior colliculus in the hedgehog tenrec. *Cereb. Cortex* 5, 338–352. doi: 10.1093/cercor/5.4.338
- Kuwabara, N., and Zook, J. M. (2000). Geniculo-collicular descending projections in the gerbil. *Brain Res.* 878, 79–87. doi: 10.1016/s0006-8993(00)02695-0
- Lam, Y.-W., and Sherman, S. M. (2010). Functional organization of the somatosensory cortical layer 6 feedback to the thalamus. *Cereb. Cortex* 20, 13–24. doi: 10.1093/cercor/bhp077
- Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* 15, 1605–1615. doi: 10.1523/JNEUROSCI.15-02-01605.1995
- Lee, C. C., and Sherman, S. M. (2010). Drivers and modulators in the central auditory pathways. *Front. Neurosci.* 4:79. doi: 10.3389/neuro.01.014.2010
- Lee, T. S., and Mumford, D. (2003). Hierarchical bayesian inference in the visual cortex. *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.* 20, 1434–1448. doi: 10.1364/josaa.20.001434
- Lee, T. S., and Nguyen, M. (2001). Dynamics of subjective contour formation in the early visual cortex. *Proc. Natl. Acad. Sci. U S A* 98, 1907–1911. doi: 10.1073/pnas.031579998
- Lee, T. S., Yang, C. F., Romero, R. D., and Mumford, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nat. Neurosci.* 5, 589–597. doi: 10.1038/nn0602-860
- León, A., Elgueta, D., Silva, M. A., Hamamé, C. M., and Délano, P. H. (2012). Auditory cortex basal activity modulates cochlear responses in chinchillas. *PLoS One* 7:e36203. doi: 10.1371/journal.pone.0036203
- Lesicko, A. M., Hristova, T., Maigner, K., and Llano, D. A. (2016). Connectional modularity of top-down and bottom-up multimodal inputs to the lateral cortex of the inferior colliculus. *J. Neurosci.* 36, 11037–11050. doi: 10.1523/JNEUROSCI.4134-15.2016
- Lesicko, A. M., and Llano, D. A. (2017). Impact of peripheral hearing loss on top-down auditory processing. *Hear. Res.* 343, 4–13. doi: 10.1016/j.heares.2016.05.018
- Lieberman, M., and Brown, M. (1986). Physiology and anatomy of single olivocochlear neurons in the cat. *Hear. Res.* 24, 17–36. doi: 10.1016/0378-5955(86)90003-1
- Lim, H. H., and Anderson, D. J. (2007). Antidromic activation reveals tonotopically organized projections from primary auditory cortex to the central nucleus of the inferior colliculus in guinea pig. *J. Neurophysiol.* 97, 1413–1427. doi: 10.1152/jn.00384.2006
- Liu, X., Yan, Y., Wang, Y., and Yan, J. (2010). Corticofugal modulation of initial neural processing of sound information from the ipsilateral ear in the mouse. *PLoS One* 5:e14038. doi: 10.1371/journal.pone.0014038
- Llano, D. (2013). Functional imaging of the thalamus in language. *Brain Lang.* 126, 62–72. doi: 10.1016/j.bandl.2012.06.004
- Llano, D. A., and Sherman, S. M. (2008). Evidence for nonreciprocal organization of the mouse auditory thalamocortical-corticothalamic projection systems. *J. Comp. Neurol.* 507, 1209–1227. doi: 10.1002/cne.21602
- Llano, D. A., and Sherman, S. M. (2009). Differences in intrinsic properties and local network connectivity of identified layer 5 and layer 6 adult mouse auditory corticothalamic neurons support a dual corticothalamic projection hypothesis. *Cereb. Cortex* 19, 2810–2826. doi: 10.1093/cercor/bhp050
- Loewy, D. H., Campbell, K. B., De Lugt, D. R., Elton, M., and Kok, A. (2000). The mismatch negativity during natural sleep: intensity deviants. *Clin. Neurophysiol.* 111, 863–872. doi: 10.1016/s1388-2457(00)00256-x
- Lorente De Nó, R. (1933). Vestibulo-ocular reflex arc. *Arch. Neuropsych.* 30, 245–291. doi: 10.1001/archneurpsyc.1933.02240140009001
- Lotto, A. J., and Klueder, K. R. (1998). General contrast effects in speech perception: effect of preceding liquid on stop consonant identification. *Percept. Psychophys.* 60, 602–619. doi: 10.3758/bf03206049
- Luo, F., Wang, Q., Kashani, A., and Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *J. Neurosci.* 28, 11615–11621. doi: 10.1523/JNEUROSCI.3972-08.2008
- Ma, X., and Suga, N. (2001). Corticofugal modulation of duration-tuned neurons in the midbrain auditory nucleus in bats. *Proc. Natl. Acad. Sci. U S A* 98, 14060–14065. doi: 10.1073/pnas.241517098
- Malmierca, M. S., Cristaudo, S., Pérez-González, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J. Neurosci.* 29, 5483–5493. doi: 10.1523/JNEUROSCI.4153-08.2009
- Malmierca, M. S., Le Beau, F. E. N., and Rees, A. (1996). The topographical organization of descending projections from the central nucleus of the inferior colliculus in guinea pig. *Hear. Res.* 93, 167–180. doi: 10.1016/0378-5955(95)00227-8
- Marian, V., Lam, T. Q., Hayakawa, S., and Dhar, S. (2018). Top-down cognitive and linguistic influences on the suppression of spontaneous otoacoustic emissions. *Front. Neurosci.* 12:378. doi: 10.3389/fnins.2018.00378
- Markovitz, C. D., Hogan, P. S., Wesen, K. A., and Lim, H. H. (2015). Pairing broadband noise with cortical stimulation induces extensive suppression of ascending sensory activity. *J. Neural Eng.* 12:026006. doi: 10.1088/1741-2560/12/2/026006

- Markovitz, C. D., Tang, T. T., and Lim, H. H. (2013). Tonotopic and localized pathways from primary auditory cortex to the central nucleus of the inferior colliculus. *Front. Neural Circuits* 7:77. doi: 10.3389/fncir.2013.00077
- Marsh, J. E., and Campbell, T. A. (2016). Processing complex sounds passing through the rostral brainstem: the new early filter model. *Front. Neurosci.* 10:136. doi: 10.3389/fnins.2016.00136
- Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science* 189, 226–228. doi: 10.1126/science.189.4198.226
- Mattys, S. L., White, L., and Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *J. Exp. Psychol. Gen.* 134, 477–500. doi: 10.1037/0096-3445.134.4.477
- May, P. J., and Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47, 66–122. doi: 10.1111/j.1469-8986.2009.00856.x
- McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0
- McGettigan, C., Eisner, F., Agnew, Z. K., Manly, T., Wisbey, D., and Scott, S. K. (2013). T'ain't what you say, it's the way that you say it—left insula and inferior frontal cortex work in interaction with superior temporal regions to control the performance of vocal impersonations. *J. Cogn. Neurosci.* 25, 1875–1886. doi: 10.1162/jocn_a_00427
- Meltzer, N. E., and Ryugo, D. L. (2006). Projections from auditory cortex to cochlear nucleus: a comparative analysis of rat and mouse. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 288, 397–408. doi: 10.1002/ar.a.20300
- Miller, G. A., Heise, G. A., and Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *J. Exp. Psychol.* 41, 329–335. doi: 10.1037/h0062491
- Mishra, S. K., and Lutman, M. E. (2014). Top-down influences of the medial olivocochlear efferent system in speech perception in noise. *PLoS One* 9:e85756. doi: 10.1371/journal.pone.0085756
- Mitani, A., Shimokouchi, M., and Nomura, S. (1983). Effects of stimulation of the primary auditory cortex upon colliculogeniculate neurons in the inferior colliculus of the cat. *Neurosci. Lett.* 42, 185–189. doi: 10.1016/0304-3940(83)90404-4
- Moore, B. C. (2012). *An Introduction to the Psychology of Hearing*. Leiden: Brill.
- Moore, D. R. (2012). Listening difficulties in children: bottom-up and top-down contributions. *J. Commun. Disord.* 45, 411–418. doi: 10.1016/j.jcomdis.2012.06.006
- Mulders, W., and Robertson, D. (2000). Evidence for direct cortical innervation of medial olivocochlear neurones in rats. *Hear. Res.* 144, 65–72. doi: 10.1016/s0378-5955(00)00046-0
- Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x
- Nakamoto, K. T., Mellott, J. G., Killius, J., Storey-Workley, M. E., Sowick, C. S., and Schofield, B. R. (2013). Ultrastructural examination of the corticocollicular pathway in the guinea pig: a study using electron microscopy, neural tracers and GABA immunocytochemistry. *Front. Neuroanat.* 7:13. doi: 10.3389/fnana.2013.00013
- Nakamoto, K. T., Shackleton, T. M., and Palmer, A. R. (2010). Responses in the inferior colliculus of the guinea pig to concurrent harmonic series and the effect of inactivation of descending controls. *J. Neurophysiol.* 103, 2050–2061. doi: 10.1152/jn.00451.2009
- Nassi, J. J., Lomber, S. G., and Born, R. T. (2013). Corticocortical feedback contributes to surround suppression in V1 of the alert primate. *J. Neurosci.* 33, 8504–8517. doi: 10.1523/JNEUROSCI.5124-12.2013
- Nelson, A., Schneider, D. M., Takatoh, J., Sakurai, K., Wang, F., and Mooney, R. (2013). A circuit for motor cortical modulation of auditory cortical activity. *J. Neurosci.* 33, 14342–14353. doi: 10.1523/JNEUROSCI.2275-13.2013
- Nelson, S. L., Kong, L., Liu, X., and Yan, J. (2015). Auditory cortex directs the input-specific remodeling of thalamus. *Hear. Res.* 328, 1–7. doi: 10.1016/j.heares.2015.06.016
- Norris, D., McQueen, J. M., and Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. *Behav. Brain Sci.* 23, 299–325. doi: 10.1017/s0140525x00003241
- Nourski, K. V., Steinschneider, M., Rhone, A. E., Kawasaki, H., Howard, M. A., and Banks, M. I. (2018). Auditory predictive coding across awareness states under anesthesia: an intracranial electrophysiology study. *J. Neurosci.* 38, 8441–8452. doi: 10.1523/JNEUROSCI.0967-18.2018
- O'Neill, J. J. (1957). Recognition of intelligibility test materials in context and isolation. *J. Speech. Hear. Disord.* 22, 87–90. doi: 10.1044/jshd.2201.87
- Obleser, J. (2014). Putting the listening brain in context. *Lang. Linguist. Compass* 8, 646–658. doi: 10.1111/lnc3.12098
- Ohl, F. W., and Scheich, H. (1997). Learning-induced dynamic receptive field changes in primary auditory cortex of the unanaesthetized Mongolian gerbil. *J. Comp. Physiol. A* 181, 685–696. doi: 10.1007/s003590050150
- Ohl, F. W., and Scheich, H. (2005). Learning-induced plasticity in animal and human auditory cortex. *Curr. Opin. Neurobiol.* 15, 470–477. doi: 10.1016/j.conb.2005.07.002
- Ojima, H. (1994). Terminal morphology and distribution of corticothalamic fibers originating from layers 5 and 6 of cat primary auditory cortex. *Cereb. Cortex* 4, 646–663. doi: 10.1093/cercor/4.6.646
- Patel, M., Sons, S., Yudinsev, G., Lesicko, A. M., Yang, L., Taha, G., et al. (2017). Anatomical characterization of subcortical descending projections to the inferior colliculus in mouse. *J. Comp. Neurol.* 525, 885–900. doi: 10.1002/cne.24106
- Perrot, X., Ryvlin, P., Isnard, J., Guénot, M., Catenox, H., Fischer, C., et al. (2006). Evidence for corticofugal modulation of peripheral auditory activity in humans. *Cereb. Cortex* 16, 941–948. doi: 10.1093/cercor/bhj035
- Pinto, S., Tremblay, P., Basirat, A., and Sato, M. (2019). The impact of when, what and how predictions on auditory speech perception. *Exp. Brain Res.* 237, 3143–3153. doi: 10.1007/s00221-019-05661-5
- Popelář, J., Šuta, D., Lindovský, J., Bureš, Z., Pysanenko, K., Chumak, T., et al. (2015). Cooling of the auditory cortex modifies neuronal activity in the inferior colliculus in rats. *Hear. Res.* 332, 7–16. doi: 10.1016/j.heares.2015.10.021
- Prasad, J. A., Carroll, B. J., and Sherman, S. M. (2020). Layer 5 corticofugal projections from diverse cortical areas: variations on a pattern of thalamic and extrathalamic targets. *J. Neurosci.* 40, 5785–5796. doi: 10.1523/JNEUROSCI.0529-20.2020
- Rao, R. P., and Ballard, D. H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Comput.* 9, 721–763. doi: 10.1162/neco.1997.9.4.721
- Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580
- Raz, A., Grady, S. M., Krause, B. M., Uhlrich, D. J., Manning, K. A., and Banks, M. I. (2014). Preferential effect of isoflurane on top-down vs. bottom-up pathways in sensory cortex. *Front. Syst. Neurosci.* 8:191. doi: 10.3389/fnsys.2014.00191
- Reichova, I., and Sherman, S. M. (2004). Somatosensory corticothalamic projections: distinguishing drivers from modulators. *J. Neurophysiol.* 92, 2185–2197. doi: 10.1152/jn.00322.2004
- Richardson, B. D., Hancock, K. E., and Caspary, D. M. (2013). Stimulus-specific adaptation in auditory thalamus of young and aged awake rats. *J. Neurophysiol.* 110, 1892–1902. doi: 10.1152/jn.00403.2013
- Robinson, B. L., Harper, N. S., and McAlpine, D. (2016). Meta-adaptation in the auditory midbrain under cortical influence. *Nat. Commun.* 7:13442. doi: 10.1038/ncomms13442
- Rockland, K. S., and Pandya, D. N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* 179, 3–20. doi: 10.1016/0006-8993(79)90485-2
- Rummell, B. P., Klee, J. L., and Sigurdsson, T. (2016). Attenuation of responses to self-generated sounds in auditory cortical neurons. *J. Neurosci.* 36, 12010–12026. doi: 10.1523/JNEUROSCI.1564-16.2016
- Ryugo, D. K., and Weinberger, N. M. (1976). Corticofugal modulation of the medial geniculate body. *Exp. Neurol.* 51, 377–391. doi: 10.1016/0014-4886(76)90262-4
- Saldaña, E. (1993). “Descending projections from the inferior colliculus to the cochlear nuclei in mammals,” in *The Mammalian Cochlear Nuclei. NATO ASI Series (Series A, Life Sciences)*, eds M. A. Merchán, J. M. Juiz, D. A. Godfrey and E. Mugnaini (Boston, MA: Springer), 153–165.
- Saldaña, E. (2015). All the way from the cortex: a review of auditory corticocollicular pathways. *Cerebellum* 14, 584–596. doi: 10.1007/s12311-015-0694-4

- Saldaña, E., Feliciano, M., and Mugnaini, E. (1996). Distribution of descending projections from primary auditory neocortex to inferior colliculus mimics the topography of intracollicular projections. *J. Comp. Neurol.* 371, 15–40. doi: 10.1002/(SICI)1096-9861(19960715)371:1<15::AID-CNE2>3.0.CO;2-O
- Samuel, A. G. (1981). Phonemic restoration: insights from a new methodology. *J. Exp. Psychol. Gen.* 110, 474–494. doi: 10.1037/0096-3445.110.4.474
- Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cogn. Psychol.* 32, 97–127. doi: 10.1006/cogp.1997.0646
- Schofield, B. R. (2001). Origins of projections from the inferior colliculus to the cochlear nucleus in guinea pigs. *J. Comp. Neurol.* 429, 206–220. doi: 10.1002/1096-9861(20000108)429:2<206::aid-cne3>3.0.co;2-x
- Schofield, B. R. (2009). Projections to the inferior colliculus from layer VI cells of auditory cortex. *Neuroscience* 159, 246–258. doi: 10.1016/j.neuroscience.2008.11.013
- Scholes, C., Palmer, A. R., and Sumner, C. J. (2015). Stream segregation in the anesthetized auditory cortex. *Hear. Res.* 328, 48–58. doi: 10.1016/j.heares.2015.07.004
- Schwiedrzik, C. M., and Freiwald, W. A. (2017). High-level prediction signals in a low-level area of the macaque face-processing hierarchy. *Neuron* 96, 89.e4–97.e4. doi: 10.1016/j.neuron.2017.09.007
- Senatorov, V., and Hu, B. (2002). Extracortical descending projections to the rat inferior colliculus. *Neuroscience* 115, 243–250. doi: 10.1016/s0306-4522(02)00316-0
- Serkov, F., Kienko, V., and Limanskaya, L. (1976). Responses of medial geniculate body neurons to auditory cortical stimulation. *Neurophysiology* 8, 3–9. doi: 10.1007/BF01065232
- Shao, Z., and Burkhalter, A. (1996). Different balance of excitation and inhibition in forward and feedback circuits of rat visual cortex. *J. Neurosci.* 16, 7353–7365. doi: 10.1523/JNEUROSCI.16-22-07353.1996
- Shastri, U., Mythri, H., and Kumar, U. A. (2014). Descending auditory pathway and identification of phonetic contrast by native listeners. *J. Acoust. Soc. Am.* 135, 896–905. doi: 10.1121/1.4861350
- Shipp, S. (2016). Neural elements for predictive coding. *Front. Psychol.* 7:1792. doi: 10.3389/fpsyg.2016.01792
- Sikkens, T., Bosman, C. A., and Olcese, U. (2019). The role of top-down modulation in shaping sensory processing across brain states: implications for consciousness. *Front. Syst. Neurosci.* 13:31. doi: 10.3389/fnsys.2019.00031
- Silva, S., Dias, C., and Castro, S. L. (2019). Domain-specific expectations in music segmentation. *Brain Sci.* 9:169. doi: 10.3390/brainsci9070169
- Skoe, E., and Kraus, N. (2010). Hearing it again and again: on-line subcortical plasticity in humans. *PLoS One* 5:e13645. doi: 10.1371/journal.pone.0013645
- Slater, B. J., Sons, S. K., Yudinsev, G., Lee, C. M., and Llano, D. A. (2019). Thalamocortical and intracortical inputs differentiate layer-specific mouse auditory corticocollicular neurons. *J. Neurosci.* 39, 256–270. doi: 10.1523/JNEUROSCI.3352-17.2018
- Slater, B. J., Willis, A. M., and Llano, D. A. (2013). Evidence for layer-specific differences in auditory corticocollicular neurons. *Neuroscience* 229, 144–154. doi: 10.1016/j.neuroscience.2012.10.053
- Smaragdis, P. J. (2001). *Redundancy Reduction for Computational Audition, a Unifying Approach*. Cambridge, MA: Massachusetts Institute of Technology.
- Smith, D., Aouad, R., and Keil, A. (2012). Cognitive task demands modulate the sensitivity of the human cochlea. *Front. Psychol.* 3:30. doi: 10.3389/fpsyg.2012.00030
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *J. Neurosci.* 32, 8443–8453. doi: 10.1523/JNEUROSCI.5069-11.2012
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2014). Top-down influences of written text on perceived clarity of degraded speech. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 186–199. doi: 10.1037/a0033206
- Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain Cogn.* 112, 92–97. doi: 10.1016/j.bandc.2015.11.003
- Srinivasan, M. V., Laughlin, S. B., and Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond. B* 216, 427–459. doi: 10.1098/rspb.1982.0085
- Srinivasan, S., Keil, A., Stratis, K., Woodruff Carr, K. L., and Smith, D. W. (2012). Effects of cross-modal selective attention on the sensory periphery: cochlear sensitivity is altered by selective attention. *Neuroscience* 223, 325–332. doi: 10.1016/j.neuroscience.2012.07.062
- Stebbing, K. A., Lesicko, A. M., and Llano, D. A. (2014). The auditory corticocollicular system: molecular and circuit-level considerations. *Hear. Res.* 314, 51–59. doi: 10.1016/j.heares.2014.05.004
- Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., et al. (2018). The predictive coding account of psychosis. *Biol. Psychiatry* 84, 634–643. doi: 10.1016/j.biopsych.2018.05.015
- Stewart, M. E., and Ota, M. (2008). Lexical effects on speech perception in individuals with “autistic” traits. *Cognition* 109, 157–162. doi: 10.1016/j.cognition.2008.07.010
- Sun, X., Jen, P. H.-S., Sun, D., and Zhang, S. (1989). Corticofugal influences on the responses of bat inferior collicular neurons to sound stimulation. *Brain Res.* 495, 1–8. doi: 10.1016/0006-8993(89)91212-2
- Sun, X., Xia, Q., Lai, C. H., Shum, D. K. Y., Chan, Y. S., and He, J. (2007). Corticofugal modulation of acoustically induced Fos expression in the rat auditory pathway. *J. Comp. Neurol.* 501, 509–525. doi: 10.1002/cne.21249
- Takayanagi, M., and Ojima, H. (2006). Microtopography of the dual corticothalamic projections originating from domains along the frequency axis of the cat primary auditory cortex. *Neuroscience* 142, 769–780. doi: 10.1016/j.neuroscience.2006.06.048
- Terrerros, G., and Déllano, P. H. (2015). Corticofugal modulation of peripheral auditory responses. *Front. Syst. Neurosci.* 9:134. doi: 10.3389/fnsys.2015.00134
- Torii, M., Hackett, T. A., Rakic, P., Levitt, P., and Polley, D. B. (2013). EphA signaling impacts development of topographic connectivity in auditory corticofugal systems. *Cereb. Cortex* 23, 775–785. doi: 10.1093/cercor/bhs066
- Treisman, A. M. (1964). Selective attention in man. *Br. Med. Bull.* 20, 12–16. doi: 10.1093/oxfordjournals.bmb.a070274
- Tseng, C. H., Mcneil, M. R., and Milenkovic, P. (1993). An investigation of attention allocation deficits in aphasia. *Brain Lang.* 45, 276–296. doi: 10.1006/brln.1993.1046
- Tzourio, N., El Massioui, F., Crivello, F., Joliot, M., Renault, B., and Mazoyer, B. (1997). Functional anatomy of human auditory attention studied with PET. *NeuroImage* 5, 63–77. doi: 10.1006/nimg.1996.0252
- Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398. doi: 10.1038/nn1032
- Usrey, W. M., and Sherman, S. M. (2019). Corticofugal circuits: communication lines from the cortex to the rest of the brain. *J. Comp. Neurol.* 527, 640–650. doi: 10.1002/cne.24423
- van Noorden, L. P. (1971). Rhythmic fission as a function of tone rate. *IPO Ann. Prog. Rep.* 6, 9–12.
- Vetter, D. E., Saldaña, E., and Mugnaini, E. (1993). Input from the inferior colliculus to medial olivocochlear neurons in the rat: a double label study with PHA-L and cholera toxin. *Hear. Res.* 70, 173–186. doi: 10.1016/0378-5955(93)90156-u
- Villa, A. E. P., Rouiller, E., Simm, G., Zurita, P., De Ribaupierre, Y., and De Ribaupierre, F. (1991). Corticofugal modulation of the information processing in the auditory thalamus of the cat. *Exp. Brain Res.* 86, 506–517. doi: 10.1007/BF00230524
- Villa, A. E. P., Tetko, I. V., Dutoit, P., De Ribaupierre, Y., and De Ribaupierre, F. (1999). Corticofugal modulation of functional connectivity within the auditory thalamus of rat, guinea pig and cat revealed by cooling deactivation. *J. Neurosci. Methods* 86, 161–178. doi: 10.1016/s0165-0270(98)00164-2
- Vila, C.-H., Williamson, R. S., Hancock, K. E., and Polley, D. B. (2019). Optimizing optogenetic stimulation protocols in auditory corticofugal neurons based on closed-loop spike feedback. *J. Neural Eng.* 16:066023. doi: 10.1088/1741-2552/ab39cf
- Von Helmholtz, H. (1867). *Handbuch Der Physiologischen Optik*. Leipzig: Voss.
- Wang, Y., Zhang, J., Zou, J., Luo, H., and Ding, N. (2019). Prior knowledge guides speech segregation in human auditory cortex. *Cereb. Cortex* 29, 1561–1571. doi: 10.1093/cercor/bhy052
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science* 167, 392–393. doi: 10.1126/science.167.3917.392
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nat. Rev. Neurosci.* 5, 279–290. doi: 10.1038/nrn1366
- Winer, J. A., Chernock, M. L., Larue, D. T., and Cheung, S. W. (2002). Descending projections to the inferior colliculus from the posterior thalamus

- and the auditory cortex in rat, cat and monkey. *Hear. Res.* 168, 181–195. doi: 10.1016/s0378-5955(02)00489-6
- Winer, J. A., Larue, D. T., Diehl, J. J., and Hefti, B. J. (1998). Auditory cortical projections to the cat inferior colliculus. *J. Comp. Neurol.* 400, 147–174.
- Wu, Y., and Yan, J. (2007). Modulation of the receptive fields of midbrain neurons elicited by thalamic electrical stimulation through corticofugal feedback. *J. Neurosci.* 27, 10651–10658. doi: 10.1523/JNEUROSCI.1320-07.2007
- Xiao, Z., and Suga, N. (2002). Modulation of cochlear hair cells by the auditory cortex in the mustached bat. *Nat. Neurosci.* 5, 57–63. doi: 10.1038/nn786
- Xiong, X. R., Liang, F., Zingg, B., Ji, X.-Y., Ibrahim, L. A., Tao, H. W., et al. (2015). Auditory cortex controls sound-driven innate defense behaviour through corticofugal projections to inferior colliculus. *Nat. Commun.* 6:7224. doi: 10.1038/ncomms8224
- Xiong, Y., Yu, Y. Q., Chan, Y. S., and He, J. (2004). Effects of cortical stimulation on auditory-responsive thalamic neurones in anaesthetized guinea pigs. *J. Physiol.* 560, 207–217. doi: 10.1113/jphysiol.2004.067686
- Yan, J., and Ehret, G. (2001). Corticofugal reorganization of the midbrain tonotopic map in mice. *Neuroreport* 12, 3313–3316. doi: 10.1097/00001756-200110290-00033
- Yan, J., and Ehret, G. (2002). Corticofugal modulation of midbrain sound processing in the house mouse. *Eur. J. Neurosci.* 16, 119–128. doi: 10.1046/j.1460-9568.2002.02046.x
- Yan, J., and Suga, N. (1996). Corticofugal modulation of time-domain processing of biosonar information in bats. *Science* 273, 1100–1103. doi: 10.1126/science.273.5278.1100
- Yan, J., Zhang, Y., and Ehret, G. (2005). Corticofugal shaping of frequency tuning curves in the central nucleus of the inferior colliculus of mice. *J. Neurophysiol.* 93, 71–83. doi: 10.1152/jn.00348.2004
- Yan, W., and Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat. Neurosci.* 1, 54–58. doi: 10.1038/255
- Yin, P., Strait, D. L., Radtke-Schuller, S., Fritz, J. B., and Shamma, S. A. (2020). Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. *Curr. Biol.* 30, 1649.e5–1663.e5. doi: 10.1016/j.cub.2020.02.047
- Yu, Y.-Q., Xiong, Y., Chan, Y.-S., and He, J. (2004). Corticofugal gating of auditory information in the thalamus: an *in vivo* intracellular recording study. *J. Neurosci.* 24, 3060–3069. doi: 10.1523/JNEUROSCI.4897-03.2004
- Yudintsev, G., Asilador, A., Coppinger, M., Nair, K., Prasad, M., and Llano, D. A. (2019). Connectional heterogeneity in the mouse auditory corticocollicular system. *BioRxiv* [Preprint]. doi: 10.1101/571711
- Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., and Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *NeuroImage* 32, 1826–1836. doi: 10.1016/j.neuroimage.2006.04.199
- Zhang, Y., Hakes, J. J., Bonfield, S. P., and Yan, J. (2005). Corticofugal feedback for auditory midbrain plasticity elicited by tones and electrical stimulation of basal forebrain in mice. *Eur. J. Neurosci.* 22, 871–879. doi: 10.1111/j.1460-9568.2005.04276.x
- Zhang, Y., and Suga, N. (2000). Modulation of responses and frequency tuning of thalamic and collicular neurons by cortical activation in mustached bats. *J. Neurophysiol.* 84, 325–333. doi: 10.1152/jn.2000.84.1.325
- Zhang, Y., Suga, N., and Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387, 900–903. doi: 10.1038/43180
- Zhang, Y., and Yan, J. (2008). Corticothalamic feedback for sound-specific plasticity of auditory thalamic neurons elicited by tones paired with basal forebrain stimulation. *Cereb. Cortex* 18, 1521–1528. doi: 10.1093/cercor/bhm188
- Zhang, Z., Liu, C.-H., Yu, Y.-Q., Fujimoto, K., Chan, Y.-S., and He, J. (2008). Corticofugal projection inhibits the auditory thalamus through the thalamic reticular nucleus. *J. Neurophysiol.* 99, 2938–2945. doi: 10.1152/jn.00002.2008
- Zhou, X., and Jen, P. H.-S. (2005). Corticofugal modulation of directional sensitivity in the midbrain of the big brown bat, *Eptesicus fuscus*. *Hear. Res.* 203, 201–215. doi: 10.1016/j.heares.2004.12.008
- Zhou, X., and Jen, P. H.-S. (2007). Corticofugal modulation of multi-parametric auditory selectivity in the midbrain of the big brown bat. *J. Neurophysiol.* 98, 2509–2516. doi: 10.1152/jn.00613.2007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Asilador and Llano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.