Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Contributions of the subcortical auditory system to predictive coding and the neural encoding of speech

Carles Escera[1,2,3,*]

Prevalent views in cognitive neuroscience have highlighted the auditory cortex (AC) as the major neuroanatomical site for auditory cognition. Yet, this view suffers from ‘cortical myopia’ as it neglects the intricate functional architecture of the subcortical auditory pathway. Here, I will review evidence indicating that key anatomical structures in the auditory hierarchy, such as the inferior colliculus and the medial geniculate body, play major roles in statistical learning and predictive processing, thus contributing to auditory perception. Furthermore, mounting evidence supports these subcortical structures as involved in the neural encoding of speech sounds, including categorical perception, and in early language acquisition when the AC is still immature. I will argue that a brain potential known as frequency-following response provides a methodological tool to map high-level cognitive operations to the human subcortical auditory system. Future studies should emphasize the precise interplay between cortical and subcortical structures in supporting auditory cognition.

Addresses
[1] Department of Clinical Psychology and Psychobiology, University of Barcelona, Spain
[2] Institute of Neurosciences, University of Barcelona, Spain
[3] Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain

Corresponding author: cescera@ub.edu
[*] Twitter account: @carles_escera, @Brainlab_UB, @UBneuroscience

Current Opinion in Behavioral Sciences 2023, **54**:101324

This review comes from a themed issue on **Subcortical Cognition**

Edited by **Michael Ullman, Will Saban** and **Karolina Janacsek**

For complete overview of the section, please refer to the article collection, “Subcortical Cognition”

Available online 30 October 2023

Received: 4 August 2023; Revised: 14 September 2023; Accepted: 25 September 2023

https://doi.org/10.1016/j.cobeha.2023.101324

Introduction

Deciphering the neural mechanisms of high-level auditory cognition is key to understand speech, to explain the joy of music and its role in organizing many aspects of our social behaviors, and to alleviate developmental disorders such as specific language impairment or dyslexia. For several historical reasons, epistemological assumptions, and methodological limitations [1,2], the prevalent cognitive neuroscience view has attributed these functions to the auditory cortex (AC) [2]. Yet, this view suffers from ‘cortical myopia’ [1], as it neglects the complex functional architecture of the subcortical auditory pathway. Compared with other sensory systems, the auditory pathway entails an intricate network of neurons arranged in five synaptic relays before reaching the cortex, which feature ipsilateral, contralateral, recurrent, feedforward, and feedback projections [3]. Among these, the inferior colliculus (IC) of the auditory midbrain, because of its sophisticated anatomical connectivity [4], is considered the analog to the primary visual cortex [5], and has been shown to play a major role in encoding statistical regularities from the acoustic input [6,7], in gating speech sounds by experience-dependent plasticity [8••], and in reading skills [9]. In turn, the medial geniculate body (MGB) in the thalamus has been implicated — through its anatomical connections to the amygdala — in voice identification and emotion content recognition in speech [10], and also in predictive coding [11] and reading disorders [12].

For their central role in driving complex computations on the acoustic input before it reaches the cortex, this review focusses on the contributions of the IC and the MGB to two prolific areas of research that involve the subcortical auditory system in high-level cognition: predictive perception and the neural encoding of speech. I highlight that a brain potential known as the frequency-following response (FFR), generated in the ascending auditory system, provides a proxy of subcortical auditory function in humans. I further suggest that the FFR can be used to test cognitive auditory function in developing individuals (i.e. neonates and infants), before the AC reaches full maturation. Several other subcortical structures have been implicated in auditory cognition, such as the cerebellum [13,14] and the basal ganglia [15], among others [2], but these are beyond the scope of this review.

Predictive coding in the subcortical auditory system

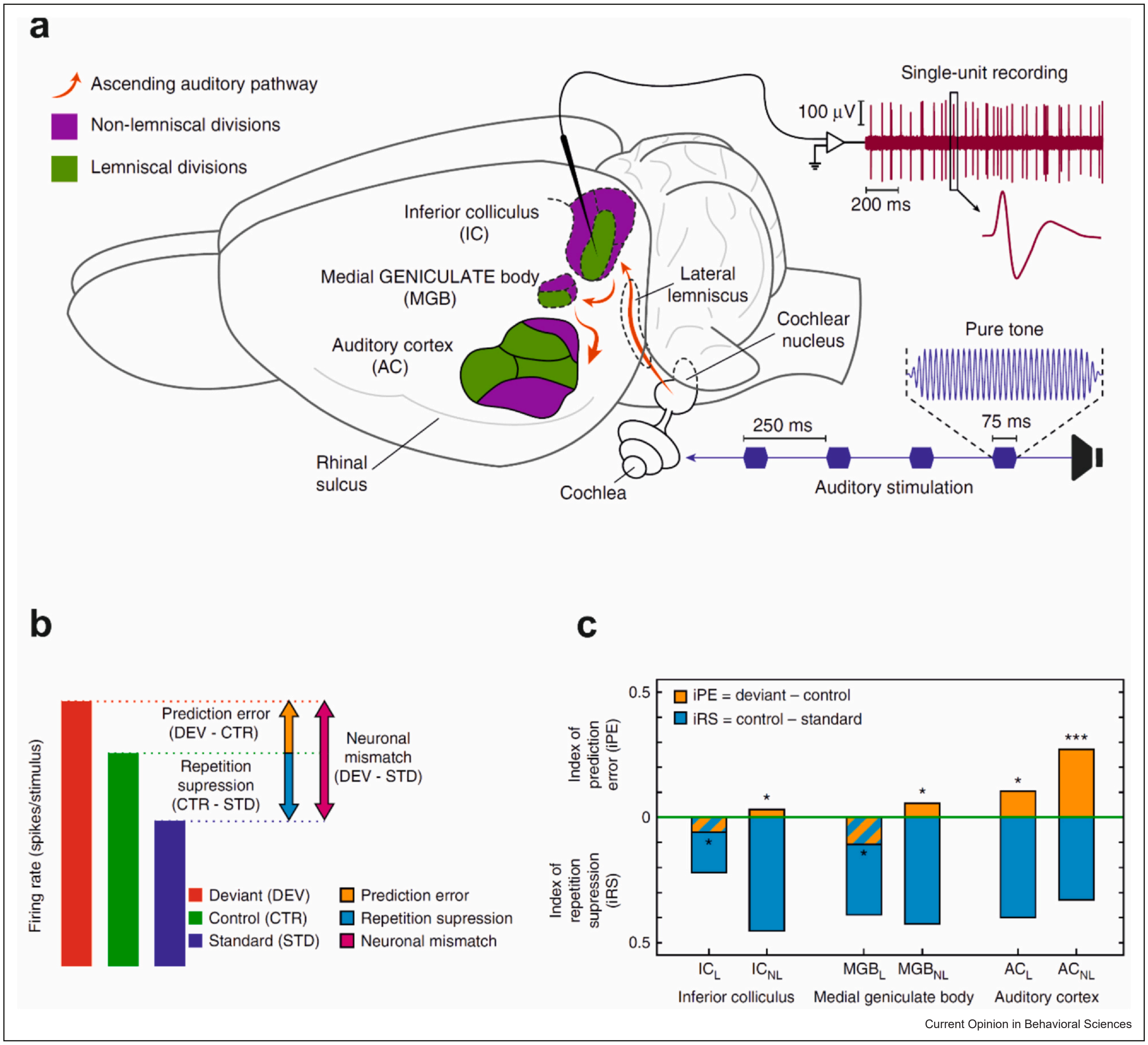
Current views of perception posit that sensory systems rely both on prior experience and contextual factors to adaptively engage with the dynamic external environment, so that neural populations modulate their responses to sound based on stimulus statistics. This is observed as suppressed or attenuated responses to repeated stimuli, whereas unexpected stimuli generate a prediction error signal [16]. According to this view, the brain is a Bayesian inference machine that computes the likelihood of sensory stimuli to occur based on previous history (i.e. predictions), with neural responses representing the mismatch between the expected event and the actual sensory input (i.e. the prediction error) [16]. In the auditory modality, evidence supporting predictive coding comes through the investigation of mismatch responses in neuronal activity in animal models and brain potentials in human studies [17,18], research that has focused mostly on cortical responses [19,20]. However, single-unit recordings in the rat have identified a subset of neurons located in nonlemniscal IC that adapted rapidly to stimulus repetition but displayed strong responses to unexpected stimulus changes [21]. Evidence supporting statistical inference in the IC was subsequently found in humans using functional magnetic resonance imaging (fMRI) [7] and brainstem evoked potentials elicited to speech [6]. Further studies extended these findings to the MGB [7,11], and demonstrated that IC/MGB neurons can encode more complex regularities than simple stimulus repetition (i.e. a violation in a tone-alternation pattern [22] or abstract rules [23]), and that prediction error emerges gradually along the auditory hierarchy, from the IC to the AC [24••] (Figure 1, see [25] for converging results in marmoset monkeys using fMRI and electrocorticography). Compelling evidence supporting that IC neurons encode for sensory predictions comes from a recent study showing neuronal responses to auditory stimuli that were actually not presented but highly predicted (i.e. omission responses) [26•].

A critical aspect in elucidating the role of the IC and MGB in auditory cognition is whether predictive responses and prediction errors are computed in these subcortical structures or otherwise inherited from the AC via the corticofugal pathway, according to the predictive coding view positing top-down propagation of predictions [16]. The corticofugal pathway is a major bundle of descending fibers that project from several auditory cortical areas to distinct portions of the MGB and IC [4], and hence may play a role in driving predictions subcortically in a top-down manner. Existing evidence is inconclusive, though. Cryogenic reversible inactivation of the AC revealed that prediction errors, at least for simple stimulus statistics (i.e. mere stimulus repetition), remained intact in both the IC [27] and the MGB [28]. However, a recent study using a similar design to [24] and optogenetic inactivation of the cortico-collicular bundle in awake mice challenged these findings [29••]. Specifically, it was found that inhibiting AC feedback decreased prediction error in IC, whereas repetition suppression remained unaffected. The results of this study align with comprehensive theoretical views suggesting that predictions in the auditory system would be generated in the AC and propagated as a cascade of generative models to the subcortical pathways via the corticofugal system [30•,31]. However, these results are in contrast with others showing statistical learning of sound patterns in IC independent of corticofugal modulation [32]. Furthermore, methodological differences (animal species vs. human participants, arousal status, and active vs. passive listening) as well as the complexity of the implicit statistics in the acoustic environment investigated, may explain these discrepancies. A recent promising approach entailing the use of two concurrent independent generative models (one based on local statistics, one driven by the task at hand) and computational modeling, suggested that both locally generated and top-down- propagated predictions modulate human neural responses (fMRI) at each level of the auditory hierarchy (IC, MGB, and AC) [33]. Yet, future research is granted to disentangle the interplay between the different levels along the auditory hierarchy in subserving predictive auditory perception.

Neural encoding of speech

Another area of research that has highlighted a major role of the subcortical auditory system to auditory cognition relates to the neural encoding of speech sounds, and its contributions to the thorough analysis of the speech signal, as well as to human communication in a broad sense. A few studies using ultra-high-field fMRI have implicated the left MGB in speech recognition, both in quiet [34] and in noisy acoustic backgrounds [35], even in specific thalamic subdivisions (i.e. the tonotopically organized primary ventral MGB) [36], as well as in reading disorders [12]. The wealth of the evidence comes, however, from studies of the human speech–-brainstem response [37], albeit the anatomical precision of these studies is limited by the intrinsic spatial resolution of the electroencephalography (EEG) approach (see the next section for a thorough discussion on the neural generators of these responses). These studies have shown the (upper) auditory brainstem to be involved in extracting pitch-relevant information [38], in mediating improved auditory discrimination in difficult listening conditions (speech in noise) [39,40], in deficient speech encoding on a number of developmental disorders, such as dyslexia [41] and autism [42], and in boosting speech processing by experience-dependent plasticity, particularly by music training [43,44] and bilingualism [45]. Of particular relevance are the studies that highlighted abnormal brainstem encoding of speech as predictors of future literacy skills [9,46]. Overall, these studies have emphasized the role of the collective auditory experiences along the lifespan [47], as well as the involvement of a cognitive–sensorimotor–reward circuitry in driving top-down plasticity within a distributed but integrated subcortico-cortical network [8••]. Given the anatomical limitations mentioned above, an open question for future research remains whether these

Figure 1



Current Opinion in Behavioral Sciences

Emergence of prediction errors along the auditory hierarchy. **(a)** Schematic illustration of the experimental setup. Isolated neurons in relays along the rat auditory neuraxis were recorded during the presentation of pure- tone sequences. Recordings were carried out in both the lemniscal (green) and nonlemniscal (purple) divisions at each station. Stimulation sequences (not shown) feature oddball conditions (a higher/lower low-probability 'deviant' tone occurring among a frequently occurring 'standard' tone), as well as several control conditions (ascending and descending 'cascade' sequences, and a 'many-standards' sequence, see [24••] for details). **(b)** Neuronal mismatch responses (deviant vs. standard, purple arrow) to the oddball sequence were decomposed using the control conditions, so that the comparison standard control isolated repetition suppression (cyan arrow), whereas the comparison deviant-control yielded the prediction error (orange arrow). **(c)** Prediction error indices (orange) and repetition suppression indices (cyan) were computed, for all hierarchical levels and anatomical divisions, by subtracting neural responses to the control from the deviant and to the standard from control, respectively, and represented as upward-positive for prediction error index (iPEs) and downward-positive for repetition suppression index (iRS). Notice, however, that for IC_L and MGB_L, prediction error yielded negative values (i.e. the control response was larger than the deviant response), and this was represented as orange bars invading the cyan area for repetition suppression. The results unveiled a gradual emergence of prediction error in the nonprimary auditory hierarchy, from IC, to MGB, and to AC, as well as in lemniscal AC, suggesting predictive coding as an intrinsic functional property of the entire auditory system.
Adapted with permission from [24••] under the Creative Commons Attribution 4.0 International License.

plastic changes pertain to the entire network or are implemented at specific levels (i.e. IC, MGB), and whether different experiences impact on different network levels and/or on the encoding of different speech attributes.

A key aspect in disentangling the role of the IC/MGB to high-level auditory cognition is whether they implement complex computational capabilities supporting categorical speech perception. Two approaches have been undertaken.

In one of such approaches, brainstem responses to speech tokens are compared from individuals with different linguistic backgrounds. For example, Mandarin speakers were found to have more robust responses than nontonal language speakers to prototypical pitch contours [48]. Another study presented a very short ambiguous acoustic cue (a voice-onset time of 2 ms followed by 10 ms of silent gap) that was perceived as a different consonant by English (/b/) or Spanish (/p/) native speakers, yielding a linguistic effect on anatomically resolved brainstem responses recorded with magnetoencephalography (MEG) [49]. An alternative approach implements a within-subjects design in which participants are presented with acoustic tokens within the continuum between two vowels. Using active — but not passive [50] — listening in conditions biasing the perceptual system, provided supporting evidence [51••]. Participants had to classify into two phonetic categories seven different tokens along the vowel continuum from /u/ to /a/, presented in random and serial (forward, backward) order conditions. It was hypothesized that serial order would warp the phonetical space, facilitating categorical encoding; accordingly, brainstem responses to the same identical acoustic token were found to be enhanced in the serial compared with the random order condition [51••].

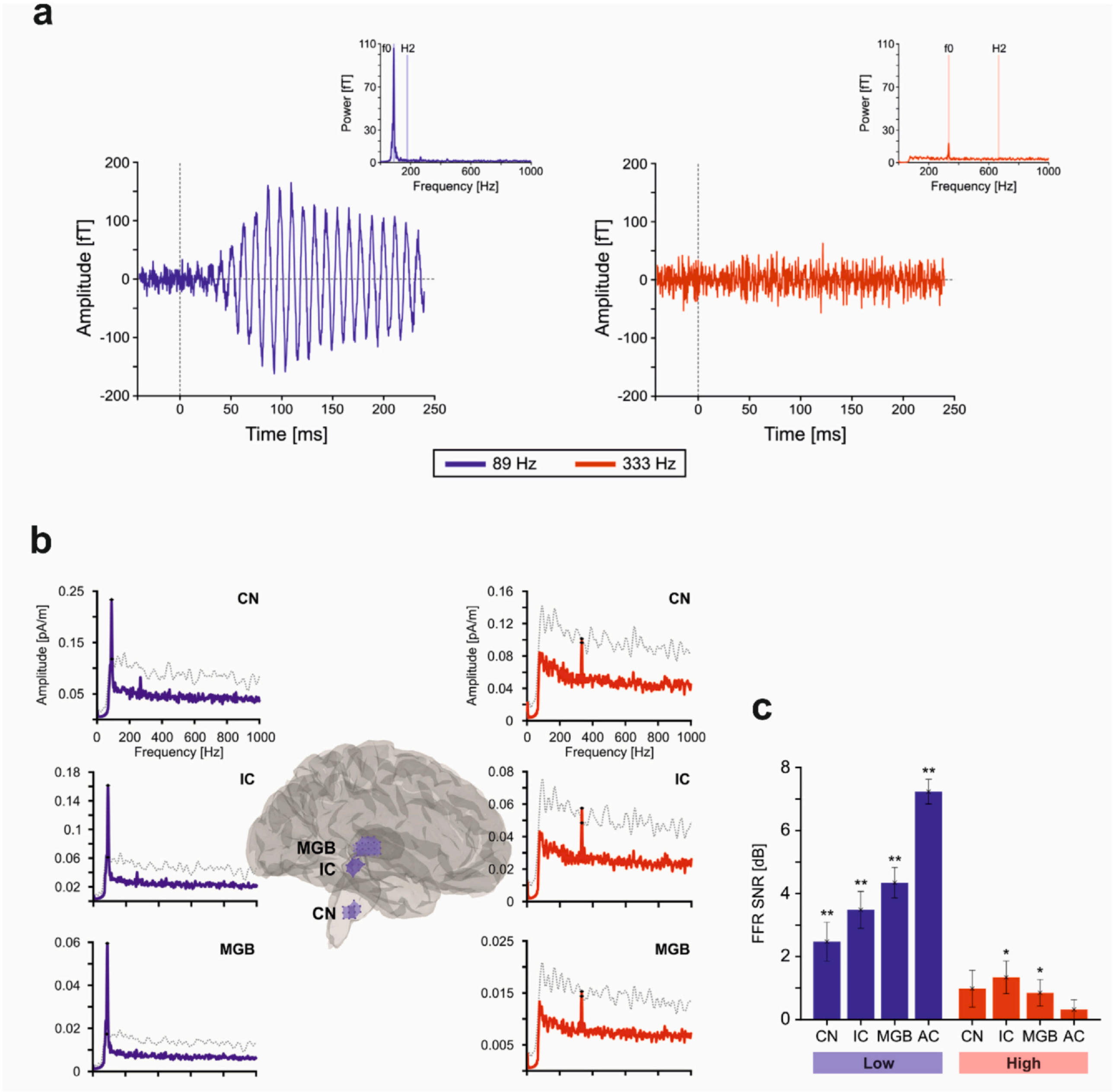
The two approaches discussed above and the results reviewed herein raise two interesting questions regarding the involvement of the subcortical auditory system in categorical speech perception. First, it seems apparent that active listening (i.e. the involvement of high-level cognition through cortical attention) is necessary for the categorical effects to be observed subcortically, hence involving the corticofugal bundles. This is further supported by studies showing enhanced source-level anatomically resolved brainstem responses (EEG) in active versus passive listening in challenging listening conditions (i.e. speech in noise [52•]), and by modulations of the brainstem response in noise according to arousal states, as indicated by cortical alpha [53] (for converging evidence involving task-modulation in gating speech recognition at the level of the MGB, see [35]). Second, studies using the across-groups approach, comparing participants with different linguistic backgrounds, support the long-term plasticity effects discussed above. In turn, studies using the within-subjects design raise the question whether the observed subcortical tuning is, in the same vein, hardwired by the linguistic experience of the participant, or induced or boosted by top-down modulation, or both. The hardwired alternative aligns with the long-term implications of the cross-linguistic studies, and supports the idea that the brainstem performs complex computations on its own, thus releasing cortical resources for higher-order functioning.

A further role of the subcortical auditory system may be in fostering early language acquisition when the entire auditory system is not yet fully mature. Despite the auditory brain being responsive to external sounds during the third trimester of pregnancy, the AC experiences a profound development during the first year of life [54], in parallel with the myelination of the auditory pathway and the acoustic radiation [55]. In these conditions, prenatal hearing [56] as well as early exposure to a rich acoustic environment right after birth may boost rapid induced plasticity in the subcortical auditory system. Recordings of speech–brainstem responses in neonates [57] have shown that infants are born with a full capacity to track voice pitch already at birth [58], that music exposure during pregnancy improves this capacity [59•], and that encoding the temporal fine structure of speech develops rapidly during the first month of life, supporting experience-dependent plasticity through acoustic exposure [60]. Whether or not this plasticity depends on corticofugal influence of an immature AC or is generated de novo by intrinsic subcortical processing awaits future investigation.

The frequency-following response as a proxy of subcortical auditory processing

The majority of the studies reviewed in the preceding section made use of scalp recordings of brainstem responses elicited to speech sounds [37]. Technically, these are called FFR, as they reflect phase-locked compound neuronal activity to the spectrotemporal components of the acoustic signal, while it is transcribed in neuronal aggregates along the auditory neuraxis. Hence, capital for the argumentation above is proving that FFRs are generated in the auditory brainstem, as suggested by seminal studies [61]. A recent influential MEG study reported a major contribution from the AC, with a right dominance [62], yet, source modeling of the scalp-recorded FFR confirmed its anatomically distributed generation, including the cochlear nucleus (CN), the IC, and the AC, and showed that the relative contribution of these sources varied as a function of sound frequency [63], owing to the synchronization cutoff of the AC at ~100 Hz [64]. Confirming evidence comes from a recent study from our group, which investigated the sources of the FFR through MEG recordings elicited to low- (89 Hz) and high- (333 Hz) tone frequencies [65•]. The results confirmed the contribution of the entire auditory hierarchy (CN, IC, MGB, and AC) to the generation of the FFR, with the AC being not significantly activated for the high-frequency tone (Figure 2). Current views support this multiple-generator hypothesis, and posit that their relative contributions may vary depending on factors such as recording techniques, stimulus features, and participant profiles [66••]. By implementing appropriate methods, such as brainstem-constrained inverse solutions (e.g. [52•]), presentation of high-frequency stimuli [65•], or analyzing neuronal activity elicited at very short latencies to

Figure 2



Current Opinion in Behavioral Sciences

The FFR as a proxy of subcortical auditory function. Distributed source modeling of the neural generators of the FFR as resolved with MEG shows neural sources along the entire auditory hierarchy, which contribute to surface responses in a frequency-dependent manner [63•]. **(a)** The time domain MEG FFR elicited to pure tones of low (89 Hz, blue) and high (333 Hz, red). The plots represent the grand-average across n = 19 participants from one single MEG channel; the upper inset shows the FFT transformation of the MEG recordings, displaying maximal spectral power at the eliciting tone frequency. **(b)** Regions of interest (ROIs) in the subcortical auditory system used to compute the amplitude spectrogram during the FFR time course (shown in blue [left] and red [right] for the low- and high-frequency tones, respectively). For statistical analysis, the peak at the frequency of interest was compared with that in the spectrogram of the baseline (i.e. the MEG activity that preceded the stimulus presentation in each trial). Sources were considered activated if the FFR peak exceeded significantly that of the baseline. (ROIs for primary AC as well as control areas are not shown). **(c)** The auditory-to-control ratio — a scalar computed to estimate the degree of activation of each source and to each frequency (see [63•] for details) — highlights the distribution of sources contributing to the generation of the surface FFR. **(d)** Of note, for the high-frequency tone (red), the cortical source was not engaged, suggesting that appropriate manipulations of stimulus frequency yield genuine subcortical FFRs.
Modified with permission from [63•].

continuous speech streams (< 10 ms, precluding cortical contribution) [67,68], I suggest that the FFR still provides a valuable proxy of subcortical auditory processing.

Conclusions and future directions

From the literature reviewed here, there is convincing evidence that both at the levels of midbrain and thalamus, complex computations occur subserving auditory cognition. Two independent areas of research support this view: predictive coding and the encoding of speech sounds, and both were the focus of this review. However, it is also currently unclear whether these computations emerge intrinsically in the subcortical nuclei, or are inherited from the AC via the corticofugal pathways. Novel experimental designs combining predictions driven by local statistics and higher-order contingencies [33], such as those related to expectations or task instructions, may help clarifying the interplay between bottom-up and top-down contributions to these processes. Furthermore, the specific contributions of the IC and the MGB to both predictive coding and speech encoding should be extended, as only few studies have addressed the two structures simultaneously. For example, the FFR described above is a compound response recruiting contributions from the entire neuraxis, but the specific roles of the collicular and thalamic relays to speech encoding remain to be clarified. Other areas of research pertaining to auditory cognition were not addressed here because their previous little interest beyond the cortex, but may provide new perspectives. For example, research on statistical learning of language, particularly of lower-level components (such as phonemes and syllables), entails both the neural encoding of speech and a predictive coding perspective [23,69–71], and may hence bring novel insights. This research has focused on cortical activity [69], with some exceptions involving the hippocampus [72], and therefore invites an examination of the potential role of the subcortical auditory system in driving statistical learning of different linguistic structures.

Several methodological developments have been put forward, such as the recording of anatomically resolved FFRs to speech [52•], the use of MEG [49,62,65•,66••], or brainstem recordings to continuous speech [67,68]. Both animal and human research addressed predictive coding, but animal research was scarce regarding the neural encoding of speech (e.g. [73]), and misses the complexity of linguistic processing carried by humans. However, animal studies have the advantage of allowing direct recording of activity in specific subdivisions of the anatomical stations of the subcortical auditory system (IC, MGB), while human EEG studies have poor spatial resolution, despite similar temporal accuracy. A potential overcoming for this latter limitation may come from the application of ultra-high-field (7 tesla or above) fMRI to map functional activation of resolved subregions of the subcortical auditory system. Although this approach has been used, with regard to auditory cognition, to map speech encoding the primary auditory thalamus [34,36], several studies have shown its potential to map the functional specialization of IC (its spatial frequency selectivity and tuning [74]), and MGB [75]. Furthermore, detailed ultra-high-field MRI 3D atlas with both activation (fMRI) and connectivity (diffusion MRI) maps has been provided [76], which can guide future research. On the other hand, animal studies allow also the direct manipulation of neural activity, while attempts to inhibit the AC in humans to test genuine contributions of IC to speech encoding were unsuccessful [77].

In summary, both human and animal research implementing novel designs combining different generative models of the sensory environment with complex sequences entailing linguistic statistics at different structural levels, appear as a promising strategy. Human research addresses the core of cognitive function, but the EEG methodology has poor spatial resolution, while MRI approaches miss the rich temporal dynamics intrinsic to the acoustic environment. In contrast, animal research is limited about questions pertaining to cognition, but may provide insights regarding the involvement of specific subdivisions of auditory stations, not yet available to human research. Together, these approaches will provide a comprehensive view of the intrinsic computations of the subcortical auditory system that contribute to auditory cognition.

Declaration of Competing Interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author is grateful to Jordi Costa-Faidella and Manuel S. Malmierca for their constructive input to early versions of this paper. This work was supported by the Spanish Ministry of Science and Innovation project PID2021-122255NB-100 (MCIN/AEI/10.13039/501100011033/FEDER, UE), the María de Maeztu Center of Excellence (CEX2021-001159-M by MCIN/AEI/10.13039/501100011033MCIN/AEI/10.13039/501100011033), the 2021SGR-0356 Excellence Research Group from the Catalan Government, and the ICREA Acadèmia Distinguished Professorship awarded to Carles Escera.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Parvizi J: **Corticocentric myopia: old bias in new cognitive sciences**. *Trends Cogn Sci* 2009, **13**:354-359, https://doi.org/10.1016/j.tics.2009.04.008

2. Janacsek K, Evans TM, Kiss M, Shah L, Blumenfeld H, Ullman MT: **Subcortical cognition: the fruit below the rind**. *Annu Rev Neurosci* 2022, **45**:361-386, https://doi.org/10.1146/annurev-neuro-110920-013544

3. Malmierca MS, Ryugo K: **Auditory system**. In *The Mouse Nervous System*. Edited by Watson C, Paxinos G, Puelles L. Elsevier; 2012:607-645.

4. Winer JA: **Decoding the auditory corticofugal systems**. *Hear Res* 2005, **207**:1-9, https://doi.org/10.1016/j.heares.2005.06.007

5. Nelken I, Fishbach A, Las L, Ulanovsky N, Farkas D: **Primary auditory cortex of cats: feature detection or something else?** *Biol Cybern* 2003, **89**:397-406, https://doi.org/10.1007/s00422-003-0445-3

6. Slabu L, Grimm S, Escera C: **Novelty detection in the human auditory brainstem**. *J Neurosci* 2012, **32**:1447-1452, https://doi.org/10.1523/jneurosci.2557-11.2012

7. Cacciaglia R, Escera C, Slabu L, Grimm S, Sanjuán A, Ventura-Campos N, Ávila C: **Involvement of the human midbrain and thalamus in auditory deviance detection**. *Neuropsychologia* 2015, **68**:51-58, https://doi.org/10.1016/j.neuropsychologia.2015.01.001

8. •• Kraus N, White-Schwoch T: **Unraveling the biology of auditory learning: a cognitive–sensorimotor–reward framework**. *Trends Cogn Sci* 2015, **19**:642-654, https://doi.org/10.1016/j.tics.2015.08.017.

A groundbreaking opinion article that proposes that high-level auditory function is sculpted by experience-dependent plasticity across a distributed but integrated network that engages cognitive, sensorimotor and reward systems.

9. White-Schwoch T, Carr KW, Thompson EC, Anderson S, Nicol T, Bradlow AR, Zecker SG, Kraus N: **Auditory processing in noise: a preschool biomarker for literacy**. *PLoS Biol* 2015, **13**:e1002196, https://doi.org/10.1371/journal.pbio.1002196

10. Pannese A, Grandjean D, Frühholz S: **Subcortical processing in auditory communication**. *Hear Res* 2015, **328**:67-77, https://doi.org/10.1016/j.heares.2015.07.003

11. Antunes FM, Nelken I, Covey E, Malmierca MS: **Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat**. *PLoS One* 2010, **5**:e14071, https://doi.org/10.1371/journal.pone.0014071

12. Díaz B, Hintz F, Kiebel SJ, Von Kriegstein K: **Dysfunction of the auditory thalamus in developmental dyslexia**. *Proc Natl Acad Sci USA* 2012, **109**:13841-13846, https://doi.org/10.1073/pnas.1119828109

13. King M, Hernandez-Castillo CR, Poldrack RA, Ivry RB, Diedrichsen J: **Functional boundaries in the human cerebellum revealed by a multi-domain task battery**. *Nat Neurosci* 2019, **22**:1371-1378, https://doi.org/10.1038/s41593-019-0436-x

14. Pinheiro A, Schwartze M, Kotz SA: **Cerebellar circuitry and auditory verbal hallucinations: an integrative synthesis and perspective**. *Neurosci Biobehav Rev* 2020, **118**:485-503, https://doi.org/10.1016/j.neubiorev.2020.08.004

15. Copland DA, Brownsett S, Iyer KK, Angwin AJ: **Corticostriatal regulation of language functions**. *Neuropsychol Rev* 2021, **31**:472-494, https://doi.org/10.1007/s11065-021-09481-9

16. Friston KJ: **The free-energy principle: a rough guide to the brain?** *Trends Cogn Sci* 2009, **13**:293-301, https://doi.org/10.1016/j.tics.2009.04.005

17. Escera C, Malmierca MS: **The auditory novelty system: an attempt to integrate human and animal research**. *Psychophysiology* 2013, **51**:111-123, https://doi.org/10.1111/psyp.12156

18. Grimm S, Escera C, Nelken I: **Early indices of deviance detection in humans and animal models**. *Biol Psychol* 2016, **116**:23-27, https://doi.org/10.1016/j.biopsycho.2015.11.017

19. Ulanovsky N, Las L, Nelken I: **Processing of low-probability sounds by cortical neurons**. *Nat Neurosci* 2003, **6**:391-398, https://doi.org/10.1038/nn1032

20. Winkler I, Denham S, Nelken I: **Modeling the auditory scene: predictive regularity representations and perceptual objects**. *Trends Cogn Sci* 2009, **13**:532-540, https://doi.org/10.1016/j.tics.2009.09.003

21. Malmierca MS, Cristaudo S, Pérez-González D, Covey E: **Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat**. *J Neurosci* 2009, **29**:5483-5493, https://doi.org/10.1523/jneurosci.4153-08.2009

22. Malmierca MS, Niño-Aguillón BE, Javier N-D, Porteros A, Pérez-González D, Escera C: **Pattern-sensitive neurons reveal encoding of complex auditory regularities in the rat inferior colliculus**. *NeuroImage* 2019, **184**:889-900, https://doi.org/10.1016/j.neuroimage.2018.10.012

23. Tabas A, Mihai G, Kiebel SJ, Trampel R, Von Kriegstein K: **Abstract rules drive adaptation in the subcortical sensory pathway**. *ELife* 2020, **9**:e64501, https://doi.org/10.7554/elife.64501

24. •• Parras GG, Javier N-D, Carbajal G, Valdés-Baizabal C, Escera C, Malmierca MS: **Neurons along the auditory pathway exhibit a hierarchical organization of prediction error**. *Nat Commun* 2017, **8**:2148, https://doi.org/10.1038/s41467-017-02038-6.

The first systematic examination of stimulus-specific adaptation and mismatch responses in single neurons across the entire auditory hierarchy (IC, MGB, AC) and auditory subdivisions (lemniscal and non-lemniscal), arousal states (anesthesia, awake preparations), and species (rat, mice).

25. Jiang Y, Komatsu M, Chen Y, Xie R, Zhang K, Xia Y, Gui P, Liang Z, Wang L: **Constructing the hierarchy of predictive auditory sequences in the marmoset brain**. *ELife* 2022, **11**:e74653, https://doi.org/10.7554/elife.74653

26. • Lao-Rodríguez AB, Przewrocki K, Pérez-González D, Alishbayli A, Yilmaz E, Malmierca MS, Englitz B: **Neuronal responses to omitted tones in the auditory brain: a neuronal correlate for predictive coding**. *Sci Adv* 2023, **9**:eabq8657, https://doi.org/10.1126/sciadv.abq8657.

Neural responses to omitted, but expected, stimuli are considered as compelling evidence that the brain operates in Bayesian mode, driving predictions about the events to occur in the environment. Omissions responses have been found in humans using EEG and fMRI, but not at single neuron level. This is the first demonstration of a single neuron response to unexpected stimulus omissions. Since stimulus repetition drove a large expectation towards the occurrence of a subsequent stimulus, the single neuron response observed to its unexpected omission strongly supports predictive processing.

27. Anderson LA, Malmierca MS: **The effect of auditory cortex deactivation on stimulus-specific adaptation in the inferior colliculus of the rat**. *Eur J Neurosci* 2012, **37**:52-62, https://doi.org/10.1111/ejn.12018

28. Antunes FM, Malmierca MS: **Effect of auditory cortex deactivation on stimulus-specific adaptation in the medial geniculate body**. *J Neurosci* 2011, **31**:17306-17316, https://doi.org/10.1523/jneurosci.1915-11.2011

29. •• Lesicko AMH, Angeloni CF, Blackwell JM, De Biasi M, Geffen MN: **Corticofugal regulation of predictive coding**. *ELife* 2022, **11**:e73289, https://doi.org/10.7554/elife.73289.

A study that nicely examines the contribution of the corticofugal pathway to predictive coding in the subcortical auditory system. The authors recorded stimulus-specific adaptation and prediction errors in the IC of the awake mouse, and inactivated the corticofugal bundle with optogenetics. Results showed that whereas stimulus-specific adaptation was unaffected by inactivation, predictions errors in IC were decreased, thus supporting the view that subcortical prediction error is driven by top-down modulation.

30. • Tabas A, Von Kriegstein K: **Adjudicating between local and global architectures of predictive processing in the subcortical auditory pathway**. *Front Neural Circuits* 2021, **15**:644743, https://doi.org/10.3389/fncir.2021.644743.

A thorough and insightful analyses of the studies investigating predictive coding in the subcortical auditory system. The authors depart from two mutually exclusive explanatory alternatives: the local architecture, supporting predictive coding to occur at low levels of the hierarchy; and the global architecture, positing that predictions are inherited subcortically from top-down propagation. After exploring single-unit animal and EEG/fMRI human evidence, the authors conclude in favor of the global architecture.

31. Asilador AR, Llano DA: **Top-down inference in the auditory system: potential roles for corticofugal projections**. *Front Neural Circuits* 2021, **14**:615259, https://doi.org/10.3389/fncir.2020.615259

32. Cruces-Solis H, Jing Z, Babaev O, Rubin JE, Gür B, Krueger-Burg D, Strenzke N, de Hoz L: **Auditory midbrain coding of statistical learning that results from discontinuous sensory stimulation**. *PLoS Biol* 2018, **16**:e2005114, https://doi.org/10.1371/journal.pbio.2005114

33. Tabas A., von Kriegstein K. : **Concurrent generative models inform prediction error in the human auditory pathway**. 2021 arXiv:2110.12252v2 [q-bio.NC]. https://doi.org/10.48550/arXiv.2110.12252.

34. Von Kriegstein K, Patterson RD, Griffiths TD: **Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition**. *Curr Biol* 2008, **18**:1855-1859, https://doi.org/10.1016/j.cub.2008.10.052

35. Mihai PG, Tschentscher N, Von Kriegstein K: **Modulation of the primary auditory thalamus when recognizing speech with background noise**. *J Neurosci* 2021, **41**:7136-7147, https://doi.org/10.1523/jneurosci.2902-20.2021

36. Mihai PG, Moerel M, de Martino F, Trampel R, Kiebel S, von Kriegstein K: **Modulation of tonotopic ventral medial geniculate body is behaviorally relevant for speech recognition**. *Elife* 2019, **8**:e44837, https://doi.org/10.7554/eLife.44837

37. Skoe E, Kraus N: **Auditory brain stem response to complex sounds: a tutorial**. *Ear Hear* 2010, **31**:302-324, https://doi.org/10.1097/aud.0b013e3181cdb272

38. Krishnan A, Gandour JT: **The role of the auditory brainstem in processing linguistically-relevant pitch patterns**. *Brain Lang* 2009, **110**:135-148, https://doi.org/10.1016/j.bandl.2009.03.005

39. Song JH, Skoe E, Banai K, Kraus N: **Training to improve hearing speech in noise: biological mechanisms**. *Cereb Cortex* 2011, **22**:1180-1190, https://doi.org/10.1093/cercor/bhr196

40. Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N: **Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance**. *J Speech Lang Hear Res* 2013, **56**:31-43, https://doi.org/10.1044/1092-4388(2012/12-0043

41. Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N: **Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia**. *Neuron* 2009, **64**:311-319, https://doi.org/10.1016/j.neuron.2009.10.006

42. Russo N, Nicol T, Trommer BL, Zecker S, Kraus N: **Brainstem transcription of speech is disrupted in children with autism spectrum disorders**. *Dev Sci* 2009, **12**:557-567, https://doi.org/10.1111/j.1467-7687.2008.00790.x

43. Wong PCM, Skoe E, Russo N, Dees TM, Kraus N: **Musical experience shapes human brainstem encoding of linguistic pitch patterns**. *Nat Neurosci* 2007, **10**:420-422, https://doi.org/10.1038/nn1872

44. Kraus N, Slater J, Thompson EC, Hornickel J, Strait DL, Nicol T, White-Schwoch T: **Music enrichment programs improve the neural encoding of speech in at-risk children**. *J Neurosci* 2014, **34**:11913-11918, https://doi.org/10.1523/jneurosci.1881-14.2014

45. Krizman J, Marian V, Shook A, Skoe E, Kraus N: **Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages**. *Proc Natl Acad Sci USA* 2012, **109**:7877-7881, https://doi.org/10.1073/pnas.1201575109

46. Neef NE, Schaadt G, Friederici AD: **Auditory brainstem responses to stop consonants predict literacy**. *Clin Neurophysiol* 2017, **128**:484-494, https://doi.org/10.1016/j.clinph.2016.12.007

47. Skoe E, Chandrasekaran B: **The layering of auditory experiences in driving experience-dependent subcortical plasticity**. *Hear Res* 2014, **311**:36-48, https://doi.org/10.1016/j.heares.2014.01.002

48. Krishnan A, Xu Y, Gandour JT, Cariani P: **Encoding of pitch in the human brainstem is sensitive to language experience**. *Cogn Brain Res* 2005, **25**:161-168, https://doi.org/10.1016/j.cogbrainres.2005.05.004

49. Zhao TC, Kuhl PK: **Linguistic effect on speech perception observed at the brainstem**. *Proc Natl Acad Sci USA* 2018, **115**:8716-8721, https://doi.org/10.1073/pnas.1800186115

50. Bidelman GM, Moreno S, Alain C: **Tracing the emergence of categorical speech perception in the human auditory system**. *NeuroImage* 2013, **79**:201-212, https://doi.org/10.1016/j.neuroimage.2013.04.093

51. •• Carter JA, Bidelman GM: **Perceptual warping exposes categorical representations for speech in human brainstem responses**. *NeuroImage* 2023, **269**:119899, https://doi.org/10.1016/j.neuroimage.2023.119899.

A study that nicely demonstrates categorical encoding of speech sounds in the subcortical auditory system, under the contribution of the AC. Participants had to classify seven different tokens along the vowel continuum from /u/ to /a/ into these wo phonetic categories. Stimuli were presented in random and serial (forward, backward) order conditions, under the hypothesis that the serial presentation would warp the phonetic space. Brainstem responses to identical acoustic tokens were enhanced in the serial compared to the random condition, hence supporting a top-down modulation of categorical subcortical encoding.

52. • Price CN, Bidelman GM: **Attention reinforces human corticofugal system to aid speech perception in noise**. *NeuroImage* 2021, **235**:118014, https://doi.org/10.1016/j.neuroimage.2021.118014.

This is groundbreaking proposal to investigate brainstem responses to speech at the source level. The authors implemented an inverse solution based on regional dipoles placed at the brainstem (i.e. IC) and bilateral primary AC, an estimated the activity within each source from the scalp recordings. By separating cortical and subcortical FFR sources this way, and implementing functional connectivity analysis, the authors showed that attention strengthened connectivity bidirectionally, but that corticofugal connectivity was disengaged during a passive listening condition.

53. Lai J, Price CN, Bidelman GM: **Brainstem speech encoding is dynamically shaped online by fluctuations in cortical α state**. *NeuroImage* 2022, **263**:119627, https://doi.org/10.1016/j.neuroimage.2022.119627

54. Moore JK, Linthicum FH: **The human auditory system: a timeline of development**. *Int J Audiol* 2007, **46**:460-478, https://doi.org/10.1080/14992020701383019

55. Zubiaurre-Elorza L, Linke AC, Herzmann C, Wild C, Duffy H, Lee DSC, Han VK, Cusack R: **Auditory structural connectivity in preterm and healthy term infants during the first postnatal year**. *Dev Psychobiol* 2018, **60**:256-264, https://doi.org/10.1002/dev.21610

56. Gervain J: **The role of prenatal experience in language development**. *Curr Opin Behav Sci* 2018, **21**:62-67, https://doi.org/10.1016/j.cobeha.2018.02.004

57. Ribas-Prats T, Almeida L, Costa-Faidella J, Plana M, Corral M, Gómez-Roig MD, Escera C: **The frequency-following response (FFR) to speech stimuli: a normative dataset in healthy newborns**. *Hear Res* 2019, **371**:28-39, https://doi.org/10.1016/j.heares.2018.11.001

58. Arenillas-Alcón S, Costa-Faidella J, Ribas-Prats T, Gómez-Roig MD, Escera C: **Neural encoding of voice pitch and formant structure at birth as revealed by frequency-following responses**. *Sci Rep* 2021, **11**:6660, https://doi.org/10.1038/s41598-021-85799-x

59. • Arenillas-Alcón S, Ribas-Prats T, Puertollano M, Mondéjar-Segovia A, Gómez-Roig MD, Costa-Faidella J, Escera C: **Prenatal daily musical exposure is associated with enhanced neural representation of speech fundamental frequency: evidence from neonatal frequency-following responses**. *Dev Sci* 2023, **26**:e13362, https://doi.org/10.1111/desc.13362.

An investigation of the effects of the prenatal exposure to music on brainstem responses to speech. The authors recorded FFRs elicited to a speech sound (/da/ for half the sample; /oa/ for the other half) right after birth in neonates, and compared responses from babies that were exposed on a daily basis to music during pregnancy (third trimester) with those from babies not exposed at this rate. Results showed enhanced responses in exposed babies, thus demonstrating that brainstem experience-dependent plasticity relevant for speech processing occurs during the fetal period.

60. Ribas-Prats T, Cordero G, Lip-Sosa D, Arenillas-Alcón S, Costa-Faidella J, Gómez-Roig MD, Escera C: **Developmental trajectory of the frequency-following response during the first six months of life**. *J Speech Lang Hear Res* 2023.

61. Chandrasekaran B, Kraus N: **The scalp-recorded brainstem response to speech: Neural origins and plasticity**.

Psychophysiology 2010, **47**:236-246, https://doi.org/10.1111/j.1469-8986.2009.00928.x

62. Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ: **Cortical contributions to the auditory frequency-following response revealed by MEG**. *Nat Commun* 2016, **7**:11070, https://doi.org/10.1038/ncomms11070

63. Bidelman GM: **Subcortical sources dominate the neuroelectric auditory frequency-following response to speech**. *NeuroImage* 2018, **175**:56-69, https://doi.org/10.1016/j.neuroimage.2018.03.060

64. Joris PX, Schreiner CE, Rees A: **Neural processing of amplitude-modulated sounds**. *Physiol Rev* 2004, **84**:541-577, https://doi.org/10.1152/physrev.00029.2003

65. • Gorina-Careta N, Kurkela JLO, Hämäläinen JA, Astikainen P, Escera C: **Neural generators of the frequency-following response elicited to stimuli of low and high frequency: a magnetoencephalographic (MEG) study**. *NeuroImage* 2021, **231**:117866, https://doi.org/10.1016/j.neuroimage.2021.117866.

This study used MEG to model the generator sources of the FFR elicited to stimuli of low (89Hz) and high (333 Hz) tone frequencies. Distributed source modelling constrained to specific ROIs within the auditory system (CN, IC, MGB, primary AC) revealed a mixture of generators of the magnetic FFR, including all these ROIs for the low frequency, with a predominance of the right cortical source. For the high frequency, however, the cortical source did not contribute significantly, thus demonstrating that the configuration of sources contributing to the generation of FFR is frequency dependent.

66. •• Coffey EBJ, Nicol T, White-Schwoch T, Chandrasekaran B, Krizman J, Skoe E, Zatorre RJ, Kraus N: **Evolving perspectives on the sources of the frequency-following response**. *Nat Commun* 2019, **10**:5036, https://doi.org/10.1038/s41467-019-13003-w.

A recent, comprehensive review on the neural contributions to the FFR. It is highlighted that FFR is a compound response entailing the entire auditory hierarchy, and that the specific source configuration may vary as a function of a number of factors, including recording method, stimulus characteristics and participant demographics.

67. Maddox RK, Lee AKC: **Auditory brainstem responses to continuous natural speech in human listeners**. *ENeuro* (5) 2018, **2018**:1-13, https://doi.org/10.1523/eneuro.0441-17 ENEURO.0441-17.2018.

68. Forte AE, Etard O, Reichenbach T: **The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention**. *ELife* 2017, **6**:e27203, https://doi.org/10.7554/elife.27203

69. Ding N, Melloni L, Zhang H, Tian X, Poppel D: **Cortical tracking of hierarchical linguistic structures in connected speech**. *Nat Neurosci* 2016, **19**:158-164, https://doi.org/10.1038/nn.4186

70. Hasson U: **The neurobiology of uncertainty: implications for statistical learning**. *Philos Trans R Soc B* 2017, **372**:20160048, https://doi.org/10.1098/rstb.2016.0048

71. Schiavo JK, Froemke RC: **Capacities and neural mechanisms for auditory statistical learning across species**. *Hear Res* 2019, **376**:97-110, https://doi.org/10.1016/j.heares.2019.02.002

72. Henin S, Turk-Browne NB, Friedman D, Liu A, Dugan P, Flinker A, Doyle W, Devinsky O, Melloni L: **Learning hierarchical sequence representations across human cortex and hippocampus**. *Sci Adv* 2021, **7**:eabc4530, https://doi.org/10.1126/sciadv.abc4530

73. He W, Ding X, Zhang R, Chen J, Zhang D, *et al*.: **Electrically-evoked frequency-following response (EFFR) in the auditory brainstem of guinea pigs**. *PLoS One* 2014, **9**:e106719, https://doi.org/10.1371/journal.pone.0106719

74. De Martino F, Moerel M, van de Moortele PF, Ugurbil K, Goebel R, Yacoub E, Formisano E: **Spatial organization of frequency preference and selectivity in the human inferior colliculus**. *Nat Commun* 2013, **4**:1386, https://doi.org/10.1038/ncomms2379

75. Moerel M, De Martino F, Uğurbil K, Yacoub E, Formisano E: **Processing of frequency and location in human subcortical auditory structures**. *Sci Rep* 2015, **5**:17048, https://doi.org/10.1038/srep17048

76. Sitek KR, Gulban OF, Calabrese E, Johnson GA, Lage-Castellanos A, Moerel M, Ghosh SS, De Martino F: **Mapping the human subcortical auditory system using histology, postmortem MRI and in vivo MRI at 7T**. *Elife* 2019, **8**:e48932, https://doi.org/10.7554/eLife.48932

77. López-Caballero F, Martin-Trias P, Ribas-Prats T, Gorina-Careta N, Bartrés-Faz D, Escera C: **Effects of cTBS on the frequency-following response and other auditory evoked potentials**. *Front Hum Neurosci* 2020, **14**:250, https://doi.org/10.3389/fnhum.2020.00250