A

α1G

► Low-Voltage-Activated Calcium Channels

α1H

► Low-Voltage-Activated Calcium Channels

α1I

► Low-Voltage-Activated Calcium Channels

Absorbing Boundary

► Decision-Making, Threshold

Accumulation of Evidence in Decision-Making

Alexander C. Huk, Leor N. Katz and Jacob L. Yates
Departments of Neuroscience and Psychology, Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA

Definition

Accumulation of evidence in decision making is the process by which noisy sensory information is sequentially sampled until sufficient evidence has accrued to favor one decision over another or others.

Detailed Description

The accumulation of evidence over time is a central topic in computational neuroscience spanning behavior, brain, and theory (Huk and Meister 2012; Shadlen et al. 2006; Usher and McClelland 2001): (1) it is a fundamental aspect of tractable forms of cognition, such as simple forms of decision making; (2) mathematical models of how

D. Jaeger, R. Jung (eds.), *Encyclopedia of Computational Neuroscience*,
https://doi.org/10.1007/978-1-0716-1006-0

evidence could (and should) be accumulated are available and have a rich history of accounting for performance in laboratory tasks; and (3) there is an apparent disconnect between the hundreds of milliseconds over which animals can integrate evidence and the individual computing elements of the brain, neurons, which integrate their inputs over a small number of milliseconds.

Although accumulating evidence over time is a central component of both cognitive function and many statistical models for decision making, the current emphasis on this topic is likely driven by the simple fact that remarkably direct neural correlates of such temporal accumulation appear to have been found in the spiking responses of neurons in brain areas such as the posterior parietal cortex, prefrontal cortex, and other areas with premotor functions (Gold and Shadlen 2007).

Mathematical Foundations

The computational neuroscience of evidence accumulation starts with Bloch's Law, a fundamental principle of sensory processing (Bloch 1885). Bloch's Law is a description of temporal summation, such that if more time results in more signal, simple behaviors such as detection will show increases in accuracy that depend on the product of stimulus intensity and time. At longer durations, a breakdown in Bloch's Law is interpreted as the limit of the temporal summation properties of the transducer. Despite its historical significance, Bloch's Law is rarely applied to modern decision-making tasks, as it focuses on "sufficiently short" durations and does not explicitly model noise, a critical tool for manipulating difficulty in many decision-making paradigms.

As opposed to the linear dependence of accuracy on time of Bloch's Law, statistical models within the "sequential sampling" framework assume that significant noise is present and that by accumulating many samples, accuracy can be improved. Under the simplest conditions of independent additive noise at each instant in time, accuracy based on ideal temporal accumulation will improve as a function of the square root of time. This root-time improvement in accuracy is directly related to the square root (n) term that many students will recognize from basic statistics.

This statistical point is the basis of a large family of models in the "sequential sampling" framework, a set of models built and adapted within statistics, physics, mathematical psychology, and neuroscience (Stone 1960; Watson 1979; Luce 1986; Link 1992; Ditterich 2006). In essence, these models assume that noisy information about a stimulus is sampled sequentially, until sufficient evidence has been accrued to favor a decision. For binary choices, the accumulation follows a noisy trajectory, where information is represented as a single quantity in which input supporting one hypothesis is evidence against another (i.e., H1 vs. H2). Mathematically, sampled evidence is weighted to support one of the two hypotheses and may be expressed as the ratio of likelihoods, the probability of observing the evidence ("e") given the hypothesis ($p(e|H1)/p(e|H2)$). The logarithm of this quantity may be summed over time to represent the degree of evidence accumulated in favor of one of the hypotheses over the other, termed the "log likelihood ratio." This value can be used to determine the optimal stopping rule for an accumulation process as implemented in the sequential probability ratio test (Wald 1947).

Evidence Accumulation in Models of Decision Making

This family of models has been extensively used to account for behavioral data and linked to psychological mechanisms: the number of samples required to reach a bound reflects decision time, and the identity of the bound reached reflects the decision. The weight of sampled evidence is a function of stimulus strength such that a strong stimulus in favor of H1 will require fewer samples to reach the H1 bound, resulting in faster reaction times and higher accuracy. For choices where H1 and H2 are equally likely, the bounds are equidistant from the starting point, the magnitude of which primarily reflects the tight coupling between decision time and accuracy: fast decisions come at the cost of accuracy, and high accuracy comes at the cost of time (the speed-accuracy

trade-off). A shift in starting point position towards H1 may reflect an a priori bias, resulting in a higher proportion of H1 choices over H2, with faster decision times.

The first major distinction between model types is whether they posit a single accumulator or multiple, racing accumulators. If there is a single accumulation process, models posit a pair of decision bounds (for a two-alternative paradigm), and the accumulation process starts in between these upper and lower limits. Thus, incoming evidence is "signed" with positive evidence pushing the accumulator towards the upper bound and negative evidence pushing the accumulator towards the lower bound. In typical "race" models, a pair of parallel and competing mechanisms each accumulates evidence in favor of their particular outcome, and the decision is determined by the first accumulator to reach its respective bound (Usher and McClelland 2001). If the signal and noise available to both accumulators are identical, such perfectly correlated racing accumulators make the same predictions as a single accumulator, despite the difference in hypothesized architecture. On the other hand, racing mechanisms can exhibit increasingly complex behaviors as additional parameters between them become statistically uncorrelated. Independent noise changes the behavior of racing versus single-variable accumulators, and more baroque models can behave quite differently than a standard single accumulator (Ditterich 2006).

The second major distinction within sequential sampling models is whether they model time as discrete or continuous. Discrete models are more straightforward to implement in computer code, while continuous models avail themselves to closed-form mathematics. While it is of course true that in the limit (i.e., as discrete time steps go to infinitesimals) discrete models become continuous, this requires some care in practice. It is critical that discrete approximations of continuous models be implemented using appropriately fine time scale steps. Otherwise, small rounding artifacts can manifest themselves in peculiar behaviors of the model. It is therefore advisable that discrete models be implemented with checks relative to known properties of the continuous-math model. Modern computers make the prospect of fine time steps less of a practical (speed) concern, but counterintuitive or quirky predictions of a discrete modeling exercise should be interpreted with caution.

The Drift-Diffusion Model

Perhaps the most common model of evidence accumulation is the (continuous time) diffusion model, increasingly referred to as the "drift-diffusion to bound" model (DDM) (Shadlen et al. 2006; Ratcliff 1978; Ratcliff and Rouder 1998; Palmer et al. 2005; Ratcliff and McKoon 2008). It stands as a hub because it is mathematically tractable (i.e., the psychometric and chronometric functions can be derived analytically) and can be flexibly extended to capture a wide range of behavioral data. The simplest diffusion models are composed of three parameters: a drift rate, bound height, and accumulator noise. The drift rate relates the stimulus units to a rate of drift of the diffusion process. The bound height is the stopping point (for two alternatives, e.g., H1 and H2, symmetrically above and below an unbiased start point). Accumulator noise is invariant across stimulus conditions and describes the noise in the accumulation process. This version of the model is overdetermined and can be rewritten as a two-parameter model with bound height or noise fixed. The diffusion model can be easily extended. For example, an additive nondecision time may be included to account for nondecisional sensory transduction and motor execution delays. Parameters added to model trial-to-trial variability in drift rate and residual time can capture phenomena like delayed reaction times for incorrect trials (Ratcliff 1978; Ratcliff and Rouder 1998; Ratcliff and McKoon 2008).

Evidence Accumulation in Neurons

The presence or implementation of diffusion-like algorithms in the brain is still a matter of significant interest and contention. The first direct evidence for a putative diffusion process in the brain

was observed in the spike rates of single neurons in the lateral intraparietal sulcus (LIP) of macaque monkeys during a stochastic motion discrimination task where the experimenter can control the amount of motion (the "evidence") on a single trial (Shadlen and Newsome 2001; Roitman and Shadlen 2002). Neurons in the LIP have a mean spike rate that ramps up for choices that result in an eye movement into their response field (RF) and ramps down for choices out of their RF. Moreover, the slope of the ramp is steeper for trials with more evidence, mirroring the drift rate of a diffusion process. Time-varying motion stimuli have further supported the notion that LIP neurons reflect temporal accumulation (integration) of relevant sensory signals (Huk and Shadlen 2005; Kiani et al. 2008). Other work has shown that the spike rates of single neurons in a range of cortical and subcortical brain areas also exhibit correlates of a diffusion process (Ratcliff et al. 2007; Ding and Gold 2012a, b).

To date, these early investigations shed little light on exactly how neurons might implement time integration, yet understanding the biophysical mechanisms may provide significant insights into the cognitive constraints of evidence accumulation (Huk and Meister 2012; Wong and Wang 2006). Likewise, the current focus on responses as averaged over neurons, trials, and sessions paints a categorical and potentially inaccurate picture of what occurs on the time scale of individual trials and decisions (Churchland et al. 2011). The next generation of work on this topic will hopefully unpack how the theoretical mechanisms of evidence accumulation are implemented by real neural circuits (Bollimunta et al. 2012).

References

Bloch AM (1885) Expériences sur la vision. C R Seance Soc Biol Paris 37:493–495

Bollimunta A, Totten D, Ditterich J (2012) Neural dynamics of choice: single-trial analysis of decision-related activity in parietal cortex. J Neurosci 32(37): 12684–12701

Churchland AK, Kiani R, Chaudhuri R, Wang XJ, Pouget A, Shadlen MN (2011) Variance as a signature of neural computations during decision making. Neuron 69(4):818–831

Ding L, Gold JI (2012a) Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. Neuron 75:865–874

Ding L, Gold JI (2012b) Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. Cereb Cortex 22:1052–1067

Ditterich J (2006) Stochastic models of decisions about motion direction: behavior and physiology. Neural Netw 19(8):981–1012

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574

Huk AC, Meister MLR (2012) Neural correlates and neural computations in posterior parietal cortex during perceptual decision-making. Front Integr Neurosci 6:86

Huk AC, Shadlen MN (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. J Neurosci 25(45):10420–10436

Kiani R, Hanks TD, Shadlen MN (2008) Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. J Neurosci 28(12):3017–3029

Link SW (1992) The wave theory of difference and similarity. Erlbaum, Hillsdale

Luce RD (1986) Response times. Oxford University Press, New York

Palmer J, Huk AC, Shadlen MN (2005) The effect of stimulus strength on the speed and accuracy of a perceptual decision. J Vis 5(5):376–404

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85(2):59–108

Ratcliff R, McKoon G (2008) The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput 20(4):873–922

Ratcliff R, Rouder JN (1998) Modeling response times for two-choice decisions. Psychol Sci 9(5):347–356

Ratcliff R, Hasegawa YT, Hasegawa RP, Smith PL, Segraves MA (2007) Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. J Neurophysiol 97(2): 1756–1774

Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area (LIP) during a combined visual discrimination reaction time task. J Neurosci 22: 9475–9489

Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol 86:1916–1936

Shadlen MN, Hanks TD, Churchland AK, Kiani R, Yang T (2006) The speed and accuracy of a simple perceptual decision: a mathematical primer. In: Doya K, Ishii S, Rao R, Pouget A (eds) Bayesian brain: probabilistic approaches to neural coding. MIT Press, Cambridge, MA, pp 209–237

Stone M (1960) Models for choice-reaction time. Psychometrika 25:251–260

Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. Psychol Rev 108(3):550–592

Wald A (1947) Sequential analysis. Wiley, New York

Watson AB (1979) Probability summation over time. Vis Res 19(5):515–522
Wong K-F, Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. J Neurosci 26(4):1314–1328

Acoustic Memory

► Auditory Memory

Acoustic Timbre Recognition

Daniel Pressnitzer[1], Trevor Agus[2] and Clara Suied[3]
[1]Département d'Études Cognitives, École Normale Supérieure, Paris, France
[2]Equipe Audition, Département d'Études Cognitives, École Normale Supérieure, Paris, France
[3]École Normale Supérieure, Paris, France

Synonyms

Auditory recognition; Sound source identification

Definition

Timbre is what allows a listener to distinguish two sounds that have otherwise the same subjective pitch, loudness, location, and duration. For instance, when orchestral musicians tune at the beginning of a concert, they all play the same note, but one can still tell the difference between instruments. This is largely because of timbre.

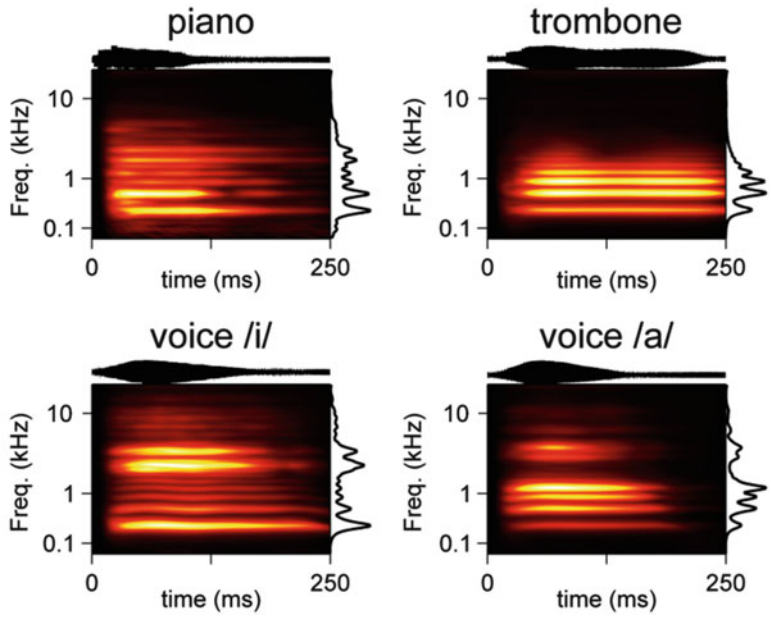
Detailed Description

The standard definition of timbre has several shortcomings. First, it says what timbre is not, rather than what it is. Second, it relates to the comparison between two sound tokens, whereas a more useful function for hearing is to associate a single timbre directly with a sound source (the timbre of the piano, the timbre of the voice of a friend). Perhaps as a consequence, there is still a lively debate about the acoustic features, mental representations, and neural mechanisms underlying timbre recognition. Here, we first outline the basic principles that make timbre such a powerful potential cue for sound source identification. Then we put forward two possible approaches to timbre, which we follow into the fields of acoustics, perception, neural mechanisms, and computational applications.

Why Do Different Sound Sources Produce Different Timbres?

Sound sources are physical objects that come in all shapes and sizes. Sound is produced when some energy makes the object vibrate. The vibrations spread around the source, which then propagate to the air and reach the ear of a listener in the form of pressure waves (Fig. 1). Simple physics shows that the wave pattern at the ear can contain a lot of information about what happened at the source (Helmholtz 1877). For instance, if the energy input was brief, such as a door knock, the chances are that the sound itself will be brief and have most of its energy concentrated around the time of the knock. After the knock, the way the door continues to vibrate is closely related to its geometry, because some wave patterns are consistent with some geometries and some are not. One such rule is that waves with low frequency and thus a long wavelength are not stable within small objects. Thus, the proportions of different frequency components that combine to make the sound of a door knock will be constrained by the size of the door. Other, more complex rules apply, depending on the shape of the object, the nature of the materials involved, and so on.

Being able to decode the intricate links between wave patterns and sound sources is extremely useful for humans and other animals. It allows the auditory system to serve as a warning sense, for instance, to identify sound-producing objects that are out of sight. For people, it is also

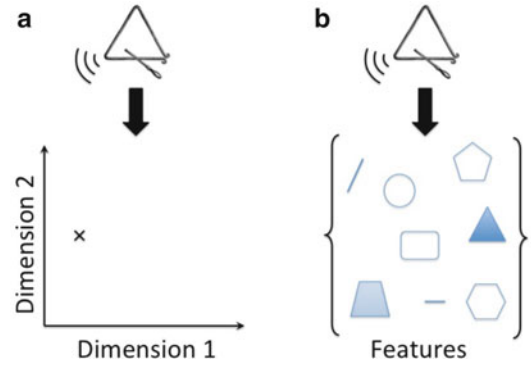


Acoustic Timbre Recognition, Fig. 1 Visual representations of four sounds with the same duration, loudness, and pitch, only differing by timbre. Each panel displays a time-frequency analysis derived from an auditory model (see Agus et al. 2012 for details). Briefly, color indicates the pattern of energy within frequency channels (*y*-axis) as it evolves over time (*x*-axis). The top trace is the corresponding pressure waveform. The right-hand trace is the average energy over time. The two instruments illustrate classic dimensions of timbre: depending on the sound source and how it is excited, the attack time can be fast (piano) or slow (trombone); the spectral center of mass can be high (piano) or low (trombone). The two vowels illustrate that other possibly more complex features may also be used to distinguish, e.g., vowels from instruments or vowels from each other

the very basis of spoken language: vowels and consonants are produced by modulating the shape of the vocal apparatus, resulting in changes in timbre that are the building blocks of oral communication.

Dimensions Versus Features

There is no consensus on what makes timbre recognition possible for human listeners. To outline current controversies, it is useful to consider two opposite viewpoints (Fig. 2). A first view is that timbre is composed of a reasonably small number of perceptual dimensions, which are subjective descriptions of sound just as pitch or loudness. Such dimensions must be metameric, in that several different sounds may project to the same point on the dimension.



Acoustic Timbre Recognition, Fig. 2 Schematic representation of the dimensions approach versus the features approach for timbre. (**a**) For the dimensions approach, all different timbres can be projected in a low-dimensional space of continuous dimensions. (**b**) For the features approach, each timbre is defined by a set of distinctive features among a very large and unordered set of possible features

A second view is that timbre recognition relies on the distinctive features of a given sound source, learned through experience and selected among a very large space of potential features. The grain of a friend's voice may be unique, which is what allows us to recognize her instantly. Such features would be conceptually different from dimensions in that a feature does not necessarily apply to all possible sound sources; in fact, it is precisely because it is unique to only a few sources (or even a single source) that it could be efficient for recognition.

It is likely that a full account of timbre will lie somewhat in between these two simplified hypotheses. However, for clarity, we continue to contrast each approach for different aspects of timbre research.

Sound Representations

To investigate timbre, it is useful to represent sound visually. Classically, this has been done with tools such as the trace of the pressure waveform over time; the spectral analysis of component frequencies through, e.g., Fourier analysis; or spectro-temporal transformations such as the short-term Fourier transform or wavelet analyses. More recently, computational models that aim to mimic peripheral or central auditory processing have been suggested (e.g., Patil et al. 2012).

In the "dimensions" approach, summary statistics are computed on sound representations to define what are referred to as descriptors of timbre. For instance, the center of mass of all frequency components of a sound produces a single number that is correlated with the apparent "brightness" of a sound (McAdams et al. 1995). In the "features" approach, the tendency is rather to maximize the richness of the representation, by including complex spectro-temporal selectivities. Such a feature-based representation need not be orderly. It can be over-complete with thousands of partially overlapping features, or sparse, in the sense that a given sound would only activate a small number of features within that large possible space (Hromadka and Zador 2009).

Perceptual Data

The basic aim of the dimensions approach is to uncover the nature and number of the perceptual dimensions underlying timbre. To this effect, statistical techniques based on multidimensional scaling have been used: a pair of sounds is presented to the listener, who has to rate how similar to each other the two sounds seem. This is repeated for all possible pairs within a given sound set. Then, the similarity judgments are treated as perceptual distances and used to obtain the dimensionality and geometry of the corresponding mental representation. For musical instruments, classic studies point toward two to three main dimensions: one related to the attack time, one related to the spectral center of mass, and one additional dimension that is less consistently observed (Grey 1977; McAdams et al. 1995). More recent investigations, using both multidimensional scaling and verbal descriptions, suggest five main dimensions with more complex interpretations (Elliott et al. 2013).

In the features approach, the focus is not on similarity but rather on the recognition of the sound source. Again, using musical instruments, fast recognition times have been observed (Agus et al. 2012), and recognition was found to be preserved even for severely impoverished signals (Suied et al. 2013). Moreover, recognition was faster and more robust for highly familiar sources such as the human voice, an observation that could not be traced back to simple acoustic dimensions (Agus et al. 2012). These results strongly suggest the existence of diagnostic features that were learned by listeners, through experience, to recognize, e.g., voices in a robust and efficient manner.

Neural Bases

Neural correlates of generic timbre dimensions have been investigated with brain imaging. Using an EEG paradigm to probe sensory memory known as mismatch negativity, it has been found that timbre dimensions such as brightness

or onset time could each be represented separately within auditory cortex (Caclin et al. 2006).

From the features perspective, single-unit recordings have uncovered a rich variety of selectivities, at many levels of the auditory system, often without any obvious ordering principle (other than by frequency). Using linear analysis techniques such as reverse correlation, spectro-temporal receptive fields have been derived. Various spectral and temporal modulation preferences have been observed, e.g., in the primary auditory cortex (Depireux et al. 2001). Adding a nonlinear component to the analysis adds another layer of complexity (Machens et al. 2004). Furthermore, the neural encoding of timbre may interact with supposedly independent sound characteristics, such as pitch or location (Bizley et al. 2009).

A further question is whether the identity of a source will be encoded by the activity of a wide network shared by many sound sources or by the activity of only a small network specifically tuned to that source category. Evidence has been put forward for both models. Using fMRI, the identity of a sound source can be inferred from distributed activity (Staeren et al. 2009). At the same time, there are clear indications of localized brain areas specialized for familiar sound sources such as the human voice (Belin 2006).

Timbre Recognition by Machines

There are several applications for acoustic timbre recognition, such as speaker identification or music information retrieval. Even though the techniques used are fast evolving and a detailed description is beyond the scope of this section, it is interesting to note that the dimensions versus features contrast can also be seen in the architectures of the computational systems.

Automatic speech recognition, which can to some extent be viewed as a timbre-decoding exercise, has a long tradition of performing classification on a small number of generic coefficients (e.g., mel-frequency cepstrum coefficients and their variants (Hermansky 1990)). For musical instruments, a descriptor-based approach has been directly inspired by the perceptual dimensions of multidimensional studies, with a reasonably small number of explicit descriptors (Peeters et al. 2011). However, other systems exist that are based on feature generation from a huge potential feature space, followed by ad hoc selection for a given classification task (Coath and Denham 2005; Pachet and Roy 2009). For musical instrument classification, machine-learning algorithms applied on a high-dimensional auditory model representation have also been successfully demonstrated (Patil et al. 2012).

Perspectives

The outstanding issues for timbre research will probably benefit from considering the various strategies available to a listener. For instance, when asked for subjective distance judgments, the most reasonable thing to do may be to abstract common dimensions to a sound set, and then use those for the comparisons. However, when asked to recognize a source as fast as possible, the mere presence of a diagnostic feature may be sufficient. The set of useful timbre dimensions or features can also depend on the task: for a same set of spoken words, different strategies are used if listeners are asked to identify the speaker or report the word content (Formisano et al. 2008). Finally, the very neural representation of timbre may be dynamically tuned to the immediate acoustic context, through rapid plasticity (Fritz et al. 2003). A fundamental reason that makes timbre so elusive may therefore be that timbre recognition is a profoundly adaptive mechanism, able to create and use opportunistic strategies that depend on the sounds and task at hand.

Cross-References

▶ Auditory Event-Related Potentials
▶ Pulse-Resonance Sounds

References

Agus TR, Suied C, Thorpe SJ, Pressnitzer D (2012) Fast recognition of musical sounds based on timbre. J Acoust Soc Am 131:4124–4133

Belin P (2006) Voice processing in human and non-human primates. Philos Trans R Soc Lond Ser B Biol Sci 361: 2091–2107

Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. J Neurosci 29:2064–2075

Caclin A, Brattico E, Tervaniemi M, Naatanen R, Morlet D, Giard MH, McAdams S (2006) Separate neural processing of timbre dimensions in auditory sensory memory. J Cognit Neurosci 18:1959–1972

Coath M, Denham SL (2005) Robust sound classification through the representation of similarity using response fields derived from stimuli during early experience. Biol Cybern 93:22–30

Depireux DA, Simon JZ, Klein DJ, Shamma SA (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. J Neurophysiol 85:1220–1234

Elliott TM, Hamilton LS, Theunissen FE (2013) Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. J Acoust Soc Am 133: 389–404

Formisano E, De Martino F, Bonte M, Goebel R (2008) "Who" is saying "what"? Brain-based decoding of human voice and speech. Science 322:970–973

Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat Neurosci 6:1216–1223

Grey JM (1977) Multidimensional perceptual scaling of musical timbres. J Acoust Soc Am 61:1270–1277

Helmholtz H (1877) On the sensations of tone. Dover, New York

Hermansky H (1990) Perceptual linear predictive (PLP) analysis of speech. J Acoust Soc Am 87:1738–1752

Hromadka T, Zador AM (2009) Representations in auditory cortex. Curr Opin Neurobiol 19:430–433

Machens CK, Wehr MS, Zador AM (2004) Linearity of cortical receptive fields measured with natural sounds. J Neurosci 24:1089–1100

McAdams S, Winsberg S, Donnadieu S, De Soete G, Krimphoff J (1995) Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. Psychol Res 58:177–192

Pachet F, Roy P (2009) Analytical features: a knowledge-based approach to audio feature generation. EURASIP J Audio Speech Music Process 2009

Patil K, Pressnitzer D, Shamma S, Elhilali M (2012) Music in our ears: the biological bases of musical timbre perception. PLoS Comput Biol 8:e1002759

Peeters G, Giordano BL, Susini P, Misdariis N, McAdams S (2011) The timbre toolbox: extracting audio descriptors from musical signals. J Acoust Soc Am 130: 2902–2916

Staeren N, Renvall H, De Martino F, Goebel R, Formisano E (2009) Sound categories are represented as distributed patterns in the human auditory cortex. Curr Biol 19:498–502

Suied C, Agus TR, Thorpe S, Pressnitzer D (2013) Processing of short auditory stimuli: the rapid audio sequential presentation paradigm (RASP). In: Moore BCJ, Patterson RD, Winter IM, Carlyon RP, Gockel HE (eds) Basic aspects of hearing: physiology and perception. Springer, New York

Action Planning

► Decision-Making, Motor Planning

Action Potential Back-Propagation

Sonia Gasparini[1] and Michele Migliore[2,3]
[1]Neuroscience Center, Louisiana State University Health Sciences Center-New Orleans, New Orleans, LA, USA
[2]Department of Neurobiology, Yale University School of Medicine, New Haven, CT, USA
[3]Institute of Biophysics, National Research Council, Palermo, Italy

Synonyms

Back-propagating action potentials (bAPs); Back-propagating spikes

Definition

Action potential (AP) back-propagation, as opposed to forward-propagation along the axon, consists of the conduction of the axonally initiated AP along neuronal dendrites, in the form of a depolarization sustained by both active and passive mechanisms. The amplitude of the depolarization generally decreases along the dendrites with increasing distance from the soma; the degree of attenuation is highly variable and depends on the neuronal type.

Detailed Description

Simultaneous recordings from dendrites, soma, and axon have shown that action potentials are generally initiated in the axon initial segment, the region with the lowest threshold for AP initiation (Stuart et al. 1997; Spruston et al. 2016). In addition to canonical forward-propagation along the axon to the presynaptic terminals, APs rapidly invade the soma and propagate back into the dendrites, where voltage-dependent channels actively support the depolarization (Spruston et al. 2016). As opposed to all-or-none axonal APs, the amplitude of back-propagating action potentials (bAPs) can be modulated and generally decreases along the dendrites with the distance from the soma. The extent of back-propagation varies widely in the different neuronal types investigated, ranging from non-decremental to almost passive (Fig. 1).

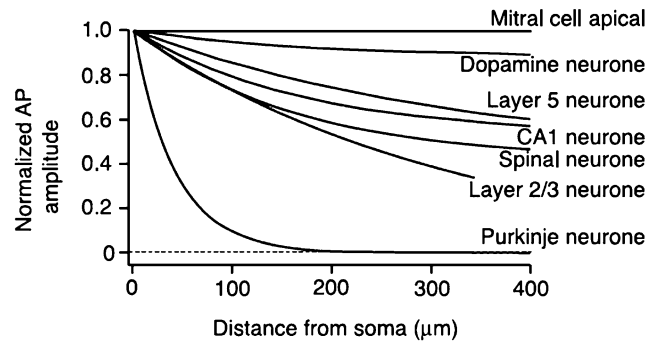
Determinants of Action Potential Back-Propagation

The main factors that determine the extent of action potential back-propagation are (a) density and properties of dendritic voltage-dependent channels, (b) neuronal morphology, and (c) neuronal activity (firing history and concurrent inputs).

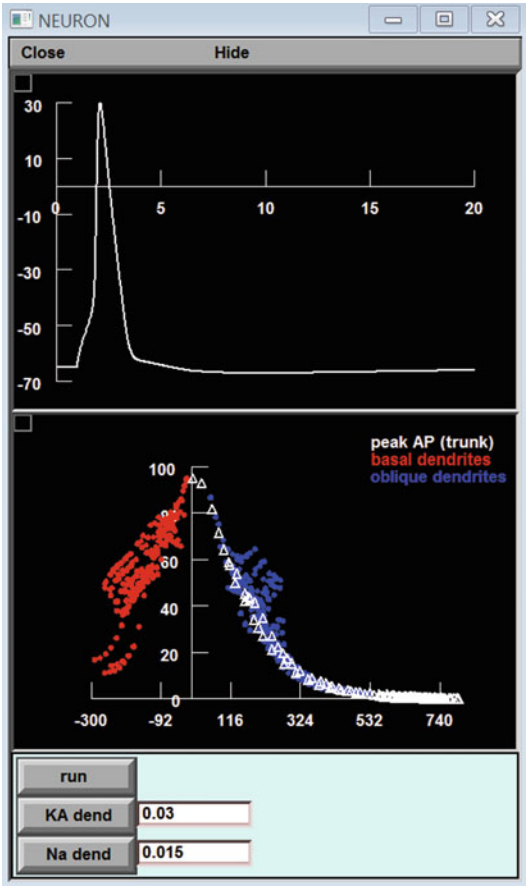
Dendritic voltage-dependent channels: Different neuronal types have distinct distributions of dendritic voltage-dependent ion channels (Migliore and Shepherd 2002, 2005; Magee 2016). AP back-propagation is sustained by dendritic voltage-dependent Na^+ channels and activates Ca^{2+} channels (Johnston et al. 1996). Dendritic K^+ channels can modulate the amplitude and extent of back-propagation; their role has been most extensively studied in hippocampal CA1 pyramidal neurons, where the density of A-type (Kv4.2) K^+ currents increases along the apical dendrites with the distance from the soma (Johnston et al. 2000). Therefore, the extent of back-propagation along the dendrites depends on the relative density of dendritic Na^+ and K^+ channels (see the interactive example in Fig. 2) and is affected by channel modulation by neurotransmitters or plasticity (Johnston et al. 1999; Magee and Johnston 2005).

Morphology: It is difficult, if not impossible, to experimentally assess the exclusive contribution of dendritic morphology, and the relation to ion channel properties, to action potential back-propagation. However, computational simulations have shown that significant differences in propagation can be attributed exclusively to dendritic geometry. When the same active and passive parameters were inserted in compartmental models with reconstructed morphologies of various neuronal types, there was a strong correlation between back-propagation and dendritic geometry (branch points and relative impedance; see Fig. 3; Vetter et al. 2001; Spruston et al. 2016). These results indicate that dendritic morphology and branching patterns have a fundamental role in shaping back-propagation.

Neuronal Activity: Dendritic excitability and AP back-propagation are highly dependent on the membrane potential and therefore on concurrent neuronal activity, as well as the firing history of neurons. Hippocampal CA1, entorhinal layer V, and neocortical layer V (but not layer II/III) pyramidal neurons show a significant



Action Potential Back-Propagation, Fig. 1 Amplitude of bAPs, normalized to the amplitude of the somatic AP, along the apical dendrite, plotted as a function of the distance from the soma, in neurons from different brain areas. (Reproduced with permission from Waters et al. 2005)



Action Potential Back-Propagation, Fig. 2 Effect of different densities of dendritic Na^+ and A-type K^+ currents on AP back-propagation in CA1 pyramidal neurons. An interactive example, where channel densities can be modified, is available at http://senselab.med.yale.edu/ModelDB/ShowModel.asp?model=148646

activity-dependent decrease in the amplitude of bAPs along the apical dendrites during high-frequency trains (Johnston et al. 1999; Gasparini 2011). In CA1 and entorhinal layer V pyramidal neurons, this behavior is mostly due to a slow, cumulative inactivation of dendritic Na^+ channels and can be reduced by neurotransmitters activating protein kinase C (Johnston et al. 1999; Magee 2016). In addition, the amplitude of bAPs at distal locations can be boosted by appropriately timed synaptic inputs (see the interactive example in Fig. 4), which promote back-propagation by inactivating A-type K^+ channels or by facilitating Na^+ channel activation (Migliore et al. 1999; Spruston 2008). On the other hand, the bAP-induced depolarization and Ca^{2+} influx can be significantly reduced by inhibition, achieved experimentally through either GABA uncaging or stimulation of individual of inhibitory interneurons (Boivin and Nedivi 2018).

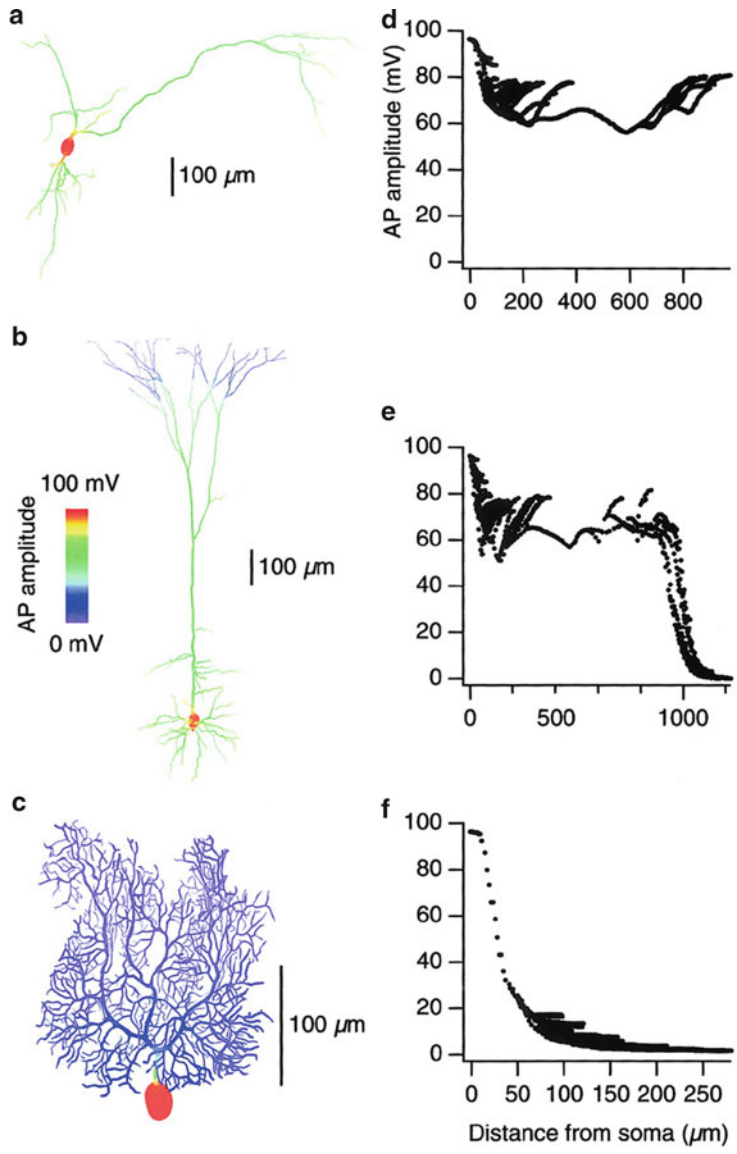
Functional Role

The depolarization and Ca^{2+} influx associated with bAPs provide feedback to the region that received the synaptic input, signaling that the neuron has generated an output. This feedback signal has important consequences on neuronal firing and synaptic plasticity. In neocortical layer V neurons, pairing bAPs with a distal synaptic input initiates a dendritic Ca^{2+} spike (back-propagation-activated Ca^{2+} or BAC spike), which results in a burst of somatic APs, possibly a marker for cortical associations (Spruston 2008; Palmer et al. 2016). In many neurons, the appropriate timing of bAPs and EPSP can cause long-term changes in synaptic efficacy (spike-timing-dependent plasticity or STDP). The depolarization associated with the bAP provides the coincidence detection needed for the removal of Mg^{2+} from NMDA glutamatergic receptors, triggering plasticity phenomena (Maheux et al. 2016). The Ca^{2+} influx caused by bAP has also been suggested to mediate dendritic release of neurotransmitters (Ludwig and Pittman 2003). The functional role of bAPs in relation to animal behaviors is more difficult to establish, since dual dendritic and somatic recordings in vivo are extremely more challenging than in vitro (but see Roome and Kuhn 2018 for recent developments); therefore dendritic signals cannot be unequivocally interpreted (Palmer et al. 2016).

Techniques to Study Back-Propagation

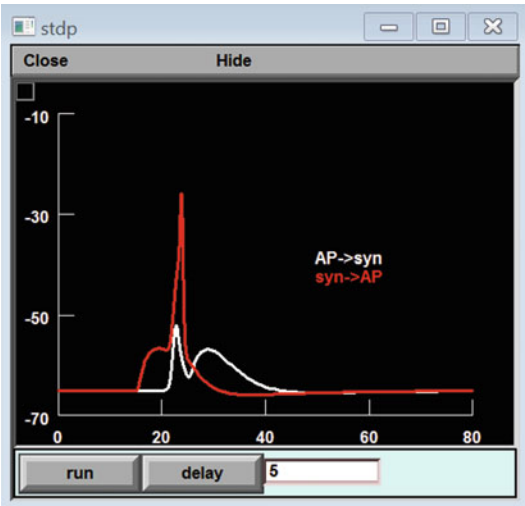
The interaction of experimental and computational methods has been essential to understanding the mechanisms underlying AP back-propagation, allowing intrinsic limitations of the individual approaches to be overcome.

Direct **electrophysiological recordings** with the patch-clamp technique in different configurations have been performed at various distances from the soma mostly on larger diameter dendrites



Action Potential Back-Propagation, Fig. 3 Effect of dendritic morphology on AP back-propagation in computational simulations with identical dendritic passive properties and channel types and densities. (**a**–**c**) color-coded representation of dendritic AP amplitude in reconstructed morphologies from a substantia nigra dopaminergic neuron (**a**), neocortical layer V neuron (**b**), and cerebellar Purkinje neuron. (**d**–**f**) bAP amplitude as a function of the distance from the soma for the cells in **a**–**c**. Reproduced with permission from Vetter et al. 2001

(Gurkiewicz and Korngreen 2006), providing information on channel distributions (Magee 2016) and back-propagation (Waters et al. 2005). **Fluorescence imaging techniques** (using Ca^{2+}- or voltage-sensitive dyes) have been used to assess smaller dendrites and compartments and/or to obtain data from many dendritic locations in a single neuron. More recently, **genetically encoded calcium and voltage indicators** have been developed and expressed in neurons, pushing the boundaries of large-scale, high-resolution monitoring of neuronal activity (Deo and Lavis 2018). These experimental techniques have advantages and limitations (Waters et al. 2005; Scanziani and Häusser 2009). For this reason, the investigation of AP back-propagation and its functional roles have greatly benefitted from **computational models** that use biophysically and morphologically accurate implementations. These models have supported, explained, and predicted



Action Potential Back-Propagation, Fig. 4 Effect of pairing a synaptic input with AP back-propagation in CA1 pyramidal neurons. If the timing is appropriate, bAP is boosted by the synaptic input (red trace). An interactive example, where the timing between AP and synaptic input can be modified, is available at http://senselab.med.yale.edu/ModelDB/ShowModel.asp?model=148646

several experimental findings on AP back-propagation (Schaefer et al. 2003; Watanabe et al. 2002). An optimal approach to studying and understanding AP back-propagation is therefore a synergistic loop, with experiments suggested by computational simulations, and experimental outcomes used to constrain the models.

Cross-References

- Action Potential Initiation
- Delayed Rectifier and A-Type Potassium Channels
- Basal Ganglia: Dopaminergic Cell Models
- High-Voltage-Activated Calcium Channels
- Long-Term Plasticity, Biophysical Models
- Low-Voltage-Activated Calcium Channels
- Multiscale Modeling of Purkinje Cells
- N-Methyl-D-Aspartate (NMDA) Receptors, Conductance Models
- Neuromodulation: Overview
- NEURON Simulation Environment
- Patch Clamp Technique
- Reduced Morphology Models
- Short-Term Plasticity, Biophysical Models
- Sodium Channels
- Spike-Timing Dependent Plasticity (STDP), Biophysical Models
- Voltage Sensitive Dye Imaging, Intrinsic Optical Signals

Acknowledgments This work was supported by the National Institutes of Health (grant NIH R01 MH115832 under the CRCNS program to SG) and by the Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2) to MM.

References

Boivin JR, Nedivi E (2018) Functional implications of inhibitory synapse placement on signal processing in pyramidal neuron dendrites. Curr Opin Neurobiol 51:16–22

Deo C, Lavis LD (2018) Synthetic and genetically encoded fluorescent neural activity indicators. Curr Opin Neurobiol 50:101–108

Gasparini S (2011) Distance- and activity-dependent modulation of spike back-propagation in layer V pyramidal neurons of the medial entorhinal cortex. J Neurophysiol 105:1372–1379

Gurkiewicz M, Korngreen A (2006) Recording, analysis, and function of dendritic voltage-gated channels. Pflugers Arch 453:283–292

Johnston D, Magee JC, Colbert CM, Cristie BR (1996) Active properties of neuronal dendrites. Annu Rev Neurosci 19:165–186

Johnston D, Hoffman DA, Colbert CM, Magee JC (1999) Regulation of back-propagating action potentials in hippocampal neurons. Curr Opin Neurobiol 9:288–292

Johnston D, Hoffman DA, Magee JC, Poolos NP, Watanabe S, Colbert CM, Migliore M (2000) Dendritic potassium channels in hippocampal pyramidal neurons. J Physiol 525(Pt 1):75–81

Ludwig M, Pittman QJ (2003) Talking back: dendritic neurotransmitter release. Trends Neurosci 26:255–261

Magee JC (2016) Voltage-gated ion channels in dendrites. In: Stuart G, Spruston N, Haüsser M (eds) Dendrites, 3rd edn. Oxford University Press, New York, pp 259–284

Magee JC, Johnston D (2005) Plasticity of dendritic function. Curr Opin Neurobiol 15:334–342

Maheux J, Froemke RC, Sjöström PJ (2016) Functional plasticity at dendritic synapses. In: Stuart G,

Spruston N, Haüsser M (eds) Dendrites, 3rd edn. Oxford University Press, New York, pp 505–555
Migliore M, Shepherd GM (2002) Emerging rules for the distributions of active dendritic conductances. Nat Rev Neurosci 3:362–370
Migliore M, Shepherd GM (2005) Opinion: an integrated approach to classifying neuronal phenotypes. Nat Rev Neurosci 6:810–818
Migliore M, Hoffman DA, Magee JC, Johnston D (1999) Role of an A-type K^+ conductance in the back-propagation of action potentials in the dendrites of hippocampal pyramidal neurons. J Comput Neurosci 7:5–15
Palmer L, Murayama M, Larkum M (2016) Dendritic integration in vitro. In: Stuart G, Spruston N, Haüsser M (eds) Dendrites, 3rd edn. Oxford University Press, New York, pp 399–427
Roome CJ, Kuhn B (2018) Simultaneous dendritic voltage and calcium imaging and somatic recording from Purkinje neurons in awake mice. Nat Commun 23:3388
Scanziani M, Häusser M (2009) Electrophysiology in the age of light. Nature 461:930–939
Schaefer AT, Larkum ME, Sakmann B, Roth A (2003) Coincidence detection in pyramidal neurons is tuned by their dendritic branching pattern. J Neurophysiol 89:3143–3154
Spruston N (2008) Pyramidal neurons: dendritic structure and synaptic integration. Nat Rev Neurosci 9:206–221
Spruston N, Stuart G, Häusser M (2016) Principles of dendritic integration. In: Stuart G, Spruston N, Haüsser M (eds) Dendrites, 3rd edn. Oxford University Press, New York, pp 351–398
Stuart G, Spruston N, Sakmann B, Häusser M (1997) Action potential initiation and backpropagation in neurons of the mammalian CNS. Trends Neurosci 20: 125–131
Vetter P, Roth A, Häusser M (2001) Propagation of action potentials in dendrites depends on dendritic morphology. J Neurophysiol 85:926–937
Watanabe S, Hoffman DA, Migliore M, Johnston D (2002) Dendritic K^+ channels contribute to spike-timing dependent long-term potentiation in hippocampal pyramidal neurons. Proc Natl Acad Sci U S A 99: 8366–8371
Waters J, Schaefer A, Sakmann B (2005) Backpropagating action potentials in neurones: measurement, mechanisms and potential functions. Prog Biophys Mol Biol 87:145–170

Further Reading

Davie JT, Kole MH, Letzkus JJ, Rancz EA, Spruston N, Stuart GJ, Häusser M (2006) Dendritic patch-clamp recording. Nat Protoc 1:1235–1247
Spruston N, Häusser M, Stuart G (2013) Information processing in dendrites and spines. In: Squire LR et al (eds) Fundamental neuroscience, 4th edn. Elsevier, Amsterdam, pp 231–260

Action Potential Initiation

Dejan Zecevic and Marko Popovic
Department of Physiology, Yale University School of Medicine, New Haven, CT, USA

Synonyms

Spike initiation

Definition

Electrical impulses (action potentials (APs) or spikes) which encode and transmit information in the nervous system are initiated in the proximal anatomical region of the axon termed axon initial segment (AIS). The voltage threshold for spike initiation and the exact location and length of the spike trigger zone (TZ) within AIS, as well as the amplitude and waveform of the action potential in different neuronal classes, depend on the geometry and passive electrical properties of a neuron as well as on the type, spatial distribution, and density of a variety of voltage-sensitive ionic channels.

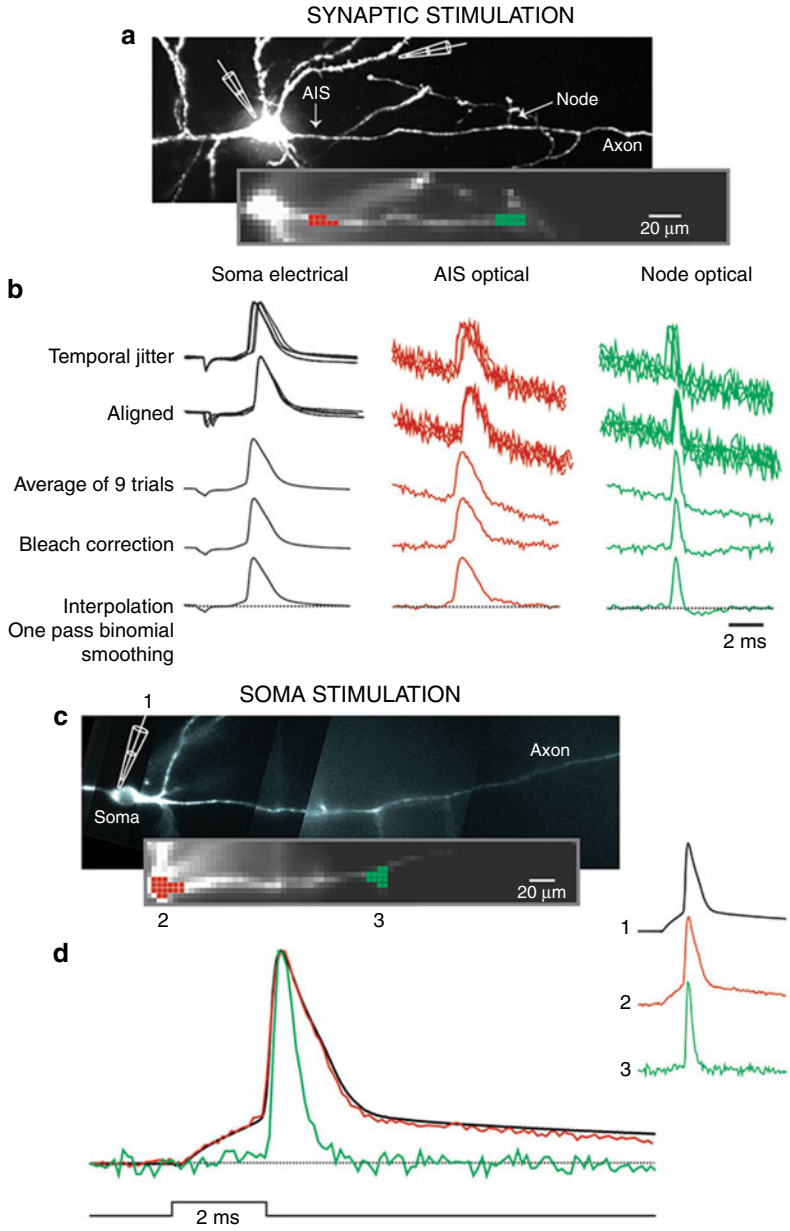
Detailed Description

There is little disagreement over attributing action potential initiation to a site in the axon initial segment (AIS) under most circumstances. The question of the exact location and length, however, of the spike trigger zone (TZ) in the axon, as defined in functional terms, is less clear. In most studies, the spike TZ was characterized by a single parameter, the distance from the soma, implying a point of initiation. The length of the initiation site, however, is fundamentally important because successful initiation and propagation of the action potential wave requires that a certain length of an axon is brought to the threshold for excitation (Rushton 1937). Besides the fundamental importance of characterizing the action potential (AP) initiation site, the spike TZ location and

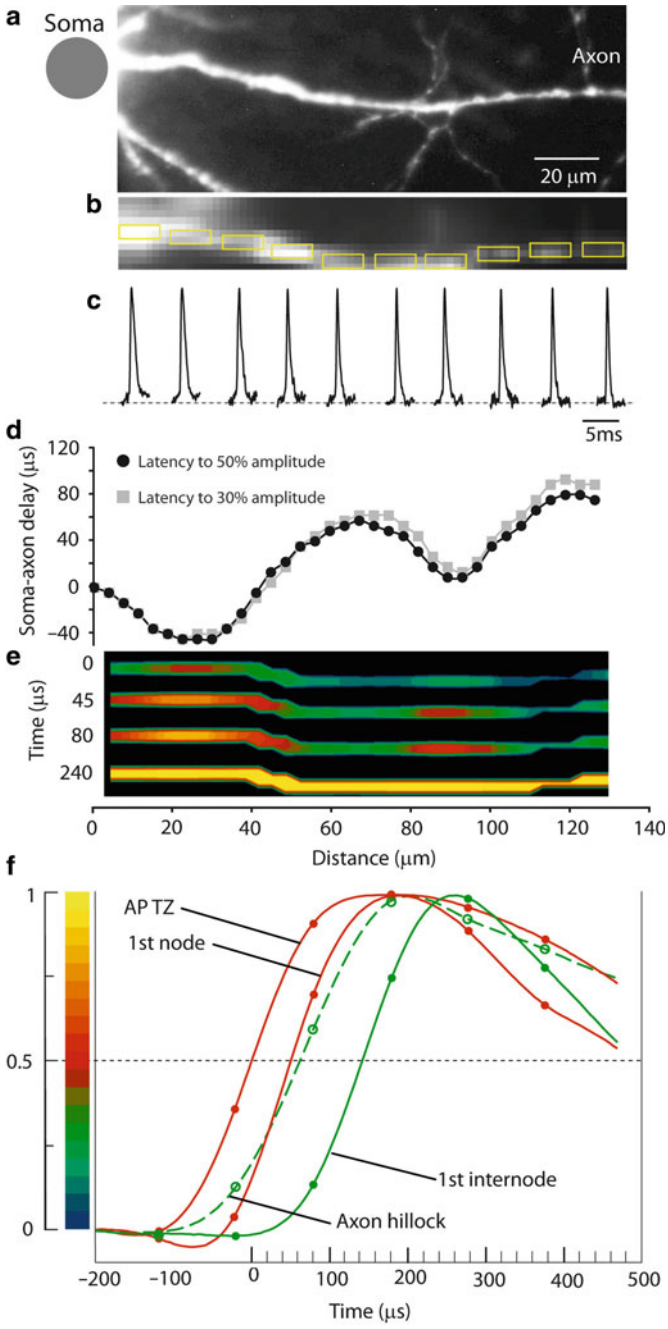
length have a recently discovered specific role in tuning neuronal computation underlying a well-defined function in auditory neurons which mediate sound source localization (Carr and Boudreau 1993; Kuba et al. 2006; Kuba and Ohmori 2009). Moreover, subsequent studies of Kuba et al. (2010) and Grubb and Burrone (2010) reported a novel finding that the structure of the spike TZ mediates an intrinsic plasticity of the axon and regulates the final stage of integration of synaptic inputs. This places a great significance on our ability to directly probe the location and length of the spike TZ under different conditions. The available information regarding TZ plasticity is based on structural data (Grubb and Burrone 2010; Kuba et al. 2010). Molecular structure of the spike TZ, however, is indirectly correlated with function in a way that is not fully understood (Fleidervish et al. 2010; Johnston 2010). Thus, the anatomical data require functional confirmation. The location and length of the spike TZ have been difficult to measure directly using electrodes because extracellular recordings cannot be interpreted with sufficient accuracy and intracellular recordings lack the necessary spatial resolution (e.g., Meeks and Mennerick 2007). Membrane potential imaging (Vm imaging) offers a unique advantage of high spatial resolution compared with electrical recordings and has been used to directly measure the location of action potential initiation in invertebrate neurons (Zecevic 1996; Antic et al. 2000) and mammalian axonal arbors (Palmer and Stuart 2006; Palmer et al. 2010). This technique was recently improved and utilized to characterize functionally relevant parameters of the spike TZ in layer 5 pyramidal neurons of the cerebral cortex (Popovic et al. 2011). The measuring technique and the analysis of data used to determine the location and length of the spike trigger zone are shown in Fig. 1.

A typical experimental measurement used to determine the location and length of the spike TZ is illustrated in Fig. 2. These two parameters are obtained directly from multisite optical recording of the membrane potential transients (Fig. 2a–c) either by investigating spike latencies at the soma/axon hillock and more distal axonal recording locations or by the inspection of the spatial distribution of membrane potential as a function of time. The soma–axon latency is plotted against recording distance from the edge of the soma in Fig. 2d. An alternative way to derive the same information from the data is to analyze a time sequence of color-coded frames showing the spatial maps of AP amplitude (Fig. 2e). The result of this analysis is a temporal series of individual frames separated by 10 μs; each showing the spatial map of membrane potential at one point in time. In Fig. 2e, four frames from this series were selected to illustrate characteristic regions along the axon that can be clearly identified during AP initiation. The red region closest to the soma was the first to cross the threshold value and reach 50% amplitude (time point 0 μs) and was identified as the AP TZ. The more distal red region appearing with a delay (45 μs time point) is likely to be the first node of Ranvier, corresponding to the issuance of an axon collateral, as indicated in the high-resolution image (Fig. 2a). The same data are shown in Fig. 2f as AP signals scaled to the same height and compared on an expanded time scale. The two red traces show AP signals from the two red areas in Fig. 2e corresponding to the spike TZ and the first node. The green dashed trace is the AP signal from the axon hillock. The green trace is the AP signal from the first internodal region. The mean length of the TZ determined from the type of data shown in Fig. 2 was 16.5 ± 1.1 μm with the mean center located at 28.9 ± 1.0 μm from the edge of the soma.

Vm imaging can be used to analyze the spatial pattern of AP propagation as revealed by monitoring transmembrane potential over longer sections of individual myelinated axons. Previously, this information was not available for any neuron. A representative experiment (well-stained neuron characterized by long axons in one plane of focus close to the surface of the slice) is illustrated in Fig. 3. The spatial plot of the soma–axon latency along an axonal section of approximately 300 μm clearly identified the position of the spike TZ and putative nodes of Ranvier; all characterized by localized reduction in soma–axon latency typical for saltatory conduction (Fig. 3a). The spatial plot of AP latency provides a functional readout for the position of the nodes of Ranvier. Figure 3c shows



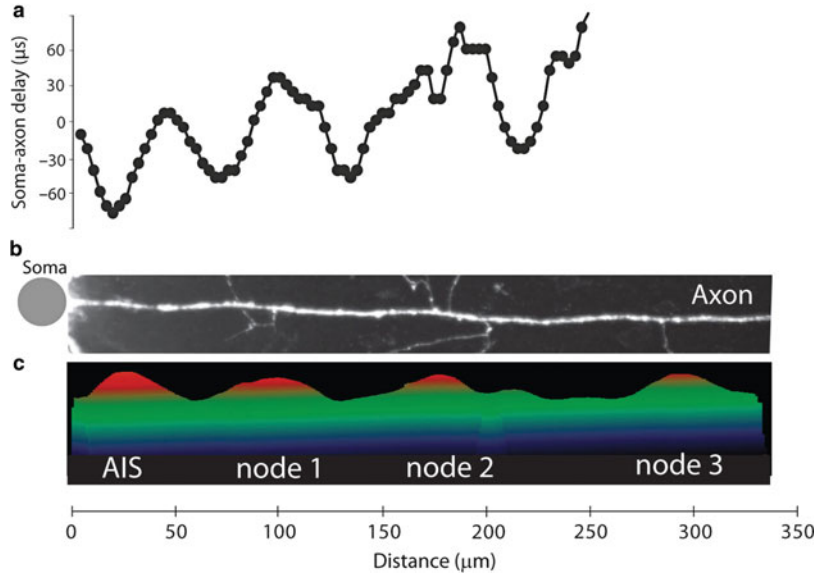
Action Potential Initiation, Fig. 1 Signal processing. (**a**) Synaptic stimulation: *Upper* image – high-resolution confocal image of a stained neuron with axon in recording position. Recording electrode attached to soma and stimulating electrode next to basal dendrite shown schematically. *Lower* image – low spatial resolution fluorescence image of the axon obtained by CCD used for V_m imaging. (**b**) Electrode recordings from soma and optical recordings from spike TZ (*red*) and from node of Ranvier (*green*). *Top traces*: raw data from nine trials showing temporal jitter in AP initiation following synaptic activation. Second row of traces: temporally aligned signals. Third row of traces: averaged signal. Fourth row of traces: bleach correction. *Bottom traces*: cubic spline interpolation with one pass of temporal smoothing. (**c**) Somatic stimulation: *Upper* image – high-resolution confocal image of another neuron with axon in recording position. *Lower* image – low spatial resolution fluorescence image of the axon obtained by CCD used for V_m imaging. Traces on *left*: AP transients from three locations: 1, electrode recording from soma; 2, optical recording from axon hillock; and 3, optical recording from the first node of Ranvier. *Bottom traces*: Superimposed signal from the same three locations



Action Potential Initiation,
Fig. 2 Measurement of the spatial distribution of membrane potential as a function of time along the proximal axon during AP initiation. (**a**) High resolution confocal image of the axon in recording position. (**b**) Low spatial resolution fluorescence image of the axon obtained by CCD used for *V*m imaging. (**c**) AP signals from 10 locations indicated by *yellow rectangles*, each 10 μm in length. (**d**) Soma–axon latency to 30% (grey) and 50% (*black*) AP amplitude as a function of distance from the cell body. The first minimum identifies the location and length of the spike *TZ*. (**e**) Time sequence of frames showing spatial profile of colour coded relative *V*m amplitude in the axon at four characteristic time points: 0 μs – *AP* initiation at *TZ*; 45 μs and 80 μs – invasion of the first node; 240 μs – peak depolarization. (**f**) Comparison of *AP* signals from four characteristic locations on an expanded time scale. The measured data points and cubic spline interpolation curves are shown. *Red* traces – *TZ* and first node; *green dashed* trace – axon hillock; *green* trace – first internodal region. Membrane potential colour scale shown on left

a color-coded spatial map of the depolarizing AP wave at one characteristic point in time which indicates the position of the spike TZ as well as nodes of Ranvier.

The voltage-sensitive dye imaging approach allows determination of the location and length of the spike TZ in the AIS of pyramidal neurons, as defined in functional terms. In addition, it is possible to characterize the AP propagation pattern in the main axon and collaterals. It is plausible to predict that this approach will facilitate the analysis of signal interactions underlying input–output



Action Potential Initiation, Fig. 3 Spatial pattern of AP initiation and propagation in an individual axon. (**a**) Soma–axon latency to 50% AP amplitude as a function of distance from the soma. (**b**) High-resolution image of an axon in recording position. (**c**) A color-coded spatial distribution of relative V_m amplitude in the axon at one characteristic point in time showing correlation between the positions of functionally determined nodes of Ranvier and axonal branch points in panel B

transformations carried out in neuronal processes of different classes of nerve cells. The question of TZ dimensions is of fundamental importance because the size of the initiation site is a critical functional parameter. Successful initiation and propagation of the action potential wave requires that a certain length of an axon is brought to threshold for excitation to generate an action current large enough to propagate. This follows from intuitive considerations, from the classical theory (Rushton 1937), as well as from more recent experimental studies (Colbert and Pan 2002; Meeks and Mennerick 2007). Additionally, Na + channel clustering in the axon, which is critical in determining the spike TZ location and length, serves an important specific function in tuning neuronal computation underlying a well-defined function (sound localization) in auditory neurons (Carr and Boudreau 1993; Kuba et al. 2006; Kuba and Ohmori 2009). These findings advocate that the size and position of the spike TZ might be cell specific in other central neurons, depending on their function. Additionally, new data (Kuba et al. 2010; Grubb and Burrone 2010) show that the structure (and, by extrapolation, the function) of the spike TZ participates in neuronal plasticity and might, in fact, be one of the key factors controlling neuronal excitability and computation (Grubb and Burrone 2010; Kuba et al. 2010; Rasband 2010).

References

Antic S, Wuskell JP, Loew L, Zecevic D (2000) Functional profile of the giant metacerebral neuron of Helix aspersa: temporal and spatial dynamics of electrical activity in situ. J Physiol 527:55–69

Carr CE, Boudreau RE (1993) An axon with a myelinated initial segment in the bird auditory system. Brain Res 628:330–334

Colbert CM, Pan E (2002) Ion channel properties underlying axonal action potential initiation in pyramidal neurons. Nat Neurosci 5:533–538

Fleidervish IA, Lasser-Ross N, Gutnick MJ, Ross WN (2010) Na+ imaging reveals little difference in action potential–evoked Na+ influx between axon and soma. Nat Neurosci 13:852–860

Grubb MS, Burrone J (2010) Activity-dependent relocation of the axon initial segment fine-tunes neuronal excitability. Nature 465:1070–1074

Johnston D (2010) The Na+ channel conundrum: axon structure versus function. Nat Neurosci 13:784–785

Kuba H, Ohmori H (2009) Roles of axonal sodium channels in precise auditory time coding at nucleus magnocellularis of the chick. J Physiol 587:87–100

Kuba H, Ishii TM, Ohmori H (2006) Axonal site of spike initiation enhances auditory coincidence detection. Nature 444:1069–1072

Kuba H, Oichi Y, Ohmori H (2010) Presynaptic activity regulates Na1 channel distribution at the axon initial segment. Nature 465:1075–1078

Meeks JP, Mennerick S (2007) Action potential initiation and propagation in CA3 pyramidal axons. J Neurophysiol 97:3460–3472

Palmer LM, Stuart GJ (2006) Site of action potential initiation in layer V pyramidal neurons. J Neurosci 26: 1854–1863

Palmer LM, Clark BA, Grundemann J, Roth A, Stuart G, Hausser M (2010) Initiation of simple and complex spikes in cerebellar Purkinje cells. J Physiol 588: 1709–1717
Popovic M, Foust AJ, McCormick DA, Zecevic D (2011) The spatio-temporal characteristics of action potential initiation in layer 5 pyramidal neurons: a voltage imaging study. J Physiol 589:4167–4187
Rasband MN (2010) The axon initial segment and the maintenance of neuronal polarity. Nat Rev Neurosci 11:552–562
Rushton WAH (1937) Initiation of the propagated disturbance. Proc R Soc Lond B Biol Sci 124:210–243
Zecevic D (1996) Multiple spike-initiation zones in single neurons revealed by voltage-sensitive dyes. Nature 381:322–325

Further Reading

Bender KJ, Trussell LO (2012) The physiology of the axon initial segment. Annu Rev Neurosci 35: 249–265
Buffington SA, Rasband MN (2011) The axon initial segment in nervous system disease and injury. Eur J Neurosci 34:1609–1619
Debanne D, Campanac E, Bialowas A, Carlier E, Alcaraz G (2011) Axon physiology. Physiol Rev 91:555–602
Kole MH, Stuart GJ (2012) Signal processing in the axon initial segment. Neuron 73:235–247
Kuba H (2012) Structural tuning and plasticity of the axon initial segment in auditory neurons. J Physiol 590: 5571–5579

Action Potential Model

► Hodgkin-Huxley Model

Action Selection

► Decision-Making, Motor Planning

Activation of the Olfactory Sensory Neurons as a Function of Odor Concentration

► Olfactory Sensory Neurons to Odor Stimuli: Mathematical Modeling of the Response

Adaptation

► Visual Aftereffects, Models of

Adaptation in Sensory Cortices, Models of

Klaus Wimmer
Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain

Synonyms

Models of pattern adaptation; Models of sensory adaptation

Definition

Models of adaptation in sensory cortices provide a functional and/or mechanistic description of the changes in neural responses and perception caused by sensory stimuli observed in the recent past. Sensory systems compute dynamic representations of the environment: cortical neurons typically adapt their "code" according to the recently received sensory input. This continuous recalibration is reflected in changes in neuronal response properties and has been interpreted as an adjustment of the limited dynamical range to compensate changes in the environment or changes in the observer. Functional models of sensory adaptation have linked these findings to optimal coding. Moreover, adaptation has also been studied in biophysical and network models, with the goal of understanding the mechanisms that give rise to adaptation in biological cortical circuits.

Detailed Description

Sensory adaptation refers to the changes in neuronal responses and perception caused by a prolonged exposure to sensory stimuli. Adaptation

is a rapid form of plasticity that has a reversible effect on neuronal selectivity: responses adapt (on short time scales) and recover to their pre-adapted state when the source of adaptation is removed. It is found ubiquitously across different sensory modalities (visual system (Kohn 2007; Clifford et al. 2007); auditory system (King et al. 2011); whisker system (Petersen et al. 2009)).

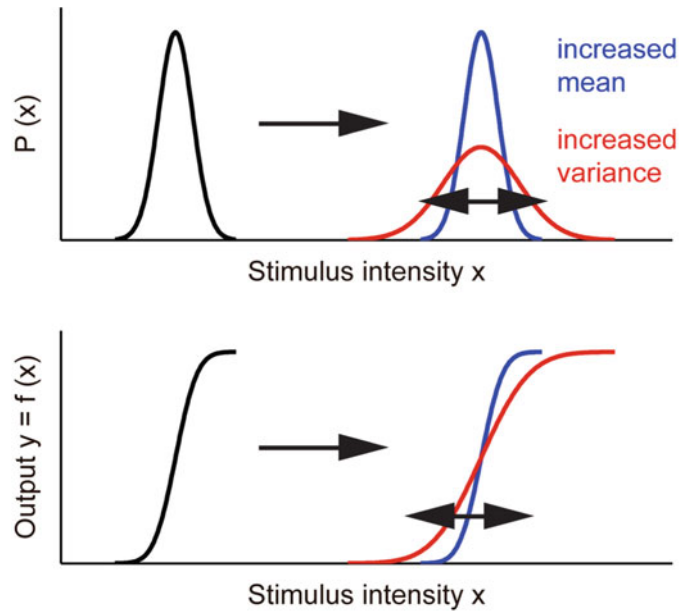
In psychophysical experiments, it can be shown that adaptation alters perception: prolonged exposure to a stimulus typically leads to "repulsive" aftereffects, which means stimuli similar to the adapting stimulus appear to be more different from the adapting stimulus than they actually are. An example is the tilt aftereffect, in which – after prolonged viewing of oblique lines – vertical lines appear briefly as if they were tilted in the opposite direction (see ▶ "Visual Aftereffects, Models of").

Here, we focus on (1) functional models of the computational principles that may underlie the adaptive encoding of sensory information and (2) models of the neuronal and synaptic mechanisms giving rise to adaptation in cortical circuits.

Adaptation as Optimal Coding

Sensory systems must encode natural stimuli which change over many orders of magnitude with the limited dynamical range of neuronal firing rates. Adaptation has been proposed to serve for adjusting the sensory representation to the current statistics of the ever changing environment (Fig. 1). This was formalized in the *efficient coding hypothesis* (grounded in information theory), which states that sensory systems seek to provide an efficient representation of the natural environment by maximizing their information transmission capacity (Barlow 1961; Laughlin 1981; Wark et al. 2007). Given the statistics of the environment, this hypothesis predicts the optimal input–output behavior of a single neuron (Fig. 1) and how it should adapt to the mean and variance of the current stimulus intensity distribution.

Experimental evidence for efficient coding has been found across a wide range of sensory modalities and species: contrast adaptation in the visual system (Kohn 2007; Clifford et al. 2007), fly visual system (Laughlin 1981; Fairhall et al. 2001), midbrain of guinea pigs (Dean et al. 2005), inferior colliculus of cats (Kvale and Schreiner 2004),



Adaptation in Sensory Cortices, Models of, Fig. 1 Adaptation to the mean and the variance of incoming sensory stimuli. According to the efficient coding hypothesis, a change in the stimulus distribution (*top*) should yield an adaptive change of a neuron's transfer function (*bottom*)

songbird auditory forebrain (Nagel and Doupe 2006), and rat barrel cortex (Maravall et al. 2007).

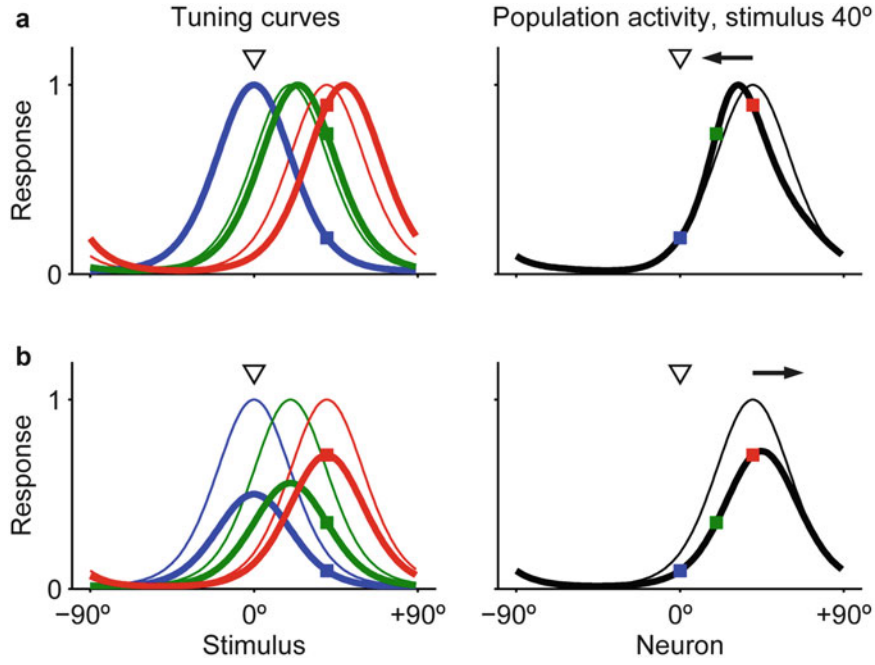
In a similar spirit, adaptation has also been linked to Bayesian inference, tackling questions such as how to optimally combine sensory observations with prior knowledge about the stimulus distribution (Doya et al. 2007). Adaptation could be interpreted as changing the prior distribution within the Bayesian theory of perception, or even the likelihood model could be affected by adapting stimuli (Stocker and Simoncelli 2006). How cortical circuits might implement Bayesian inference is an active research topic.

Modeling Underlying Neural Mechanisms

A different class of models aims at understanding the neuronal mechanisms that underlie adaptation. In some cases the mechanisms have been studied in detail experimentally, an example is contrast adaptation in the visual system. Adaptation shifts the contrast response curve of individual neurons toward higher contrast levels, and this can be well explained by the hyperpolarization of individual neurons caused by the activation of intrinsic channels (Sanchez-Vives et al. 2000). This mechanism accounts for a decrease in a neuron's firing rate due to sustained stimulation and is usually called spike-frequency adaptation; for a phenomenological model, see Benda and Herz (2003).

Often, however, the mechanisms underlying adaptation are not well characterized experimentally. Models can serve to integrate existing data and to create testable predictions in order to discern different potential mechanisms. An example is orientation adaptation, in which the prolonged exposure to a visual stimulus of a particular orientation yields changes of individual neuronal tuning curves in the early visual system (Dragoi et al. 2000) (see Fig. 2). A potential mechanism underlying this adaptation phenomenon, in addition to spike-frequency adaptation, is short-term synaptic depression. This form of rapid and reversible plasticity leads to a decrease of synaptic efficacy lasting from milliseconds to several seconds, due to prolonged presynaptic activity. This timescale is comparable to the timescale of orientation adaptation, but synaptic depression has mostly been studied in vitro, and direct experimental evidence demonstrating that it is involved in orientation adaptation is lacking. Synaptic depression can be described by a change in transmitter release probability in a phenomenological model (Abbott et al. 1997; Tsodyks and Markram 1997). Computational studies have investigated how synaptic depression, spike-frequency adaptation, and other mechanisms contribute to adaptive changes of visual cortical neurons. The different studies do however reach different conclusions about which particular intrinsic or circuit mechanism is the most likely origin of orientation adaptation (Bednar and Miikkulainen 2000; Chelaru and Dragoi 2008; Cortes et al. 2011). We argue that for a full understanding of adaptation mechanisms in early visual areas, it is necessary to study how adaptation interplays with the dynamics of local circuits, which are dominated by strong, approximately balanced, excitation and inhibition (Stimberg et al. 2009).

On the other hand, computational network models have elucidated the link between orientation adaptation and the psychophysically observed tilt aftereffect, which has long been a challenge (Teich and Qian 2003). Computational models showed how adaptive changes in the location, the width, and the magnitude of single neuron tuning curves give rise to characteristic changes in population responses (Clifford et al. 2000; Compte and Wang 2006). In fact, the relationship between changes in experimentally measured single neuron tuning curves and the response of a population of neurons to a single stimulus (which is thought to underlie perception) can be counterintuitive (Fig. 2). For example, when neuronal tuning curves shift *away* from the adapting orientation (as typically observed experimentally), the corresponding population activity shifts *toward* the adapting orientation (Fig. 2a). If, on the other hand, adaptation only causes a suppression of neuronal responses close to the adapting stimulus (Fig. 2b), the result is a shift of the population activity *away* from the adapting stimulus (as observed in the tilt aftereffect). Considered together, the combination of adaptive changes



Adaptation in Sensory Cortices, Models of, Fig. 2 Example of the relationship between adaptive changes in single neuron tuning curves and changes in the population activity, which is thought to underlie perception. (**a**) Adaptation shifts response curves of individual neurons away from the adapting stimulus. Shown are the tuning curves (responses of a neuron to visual stimuli of different orientation) of three neurons with preferred orientation 0°, 20°, and 40° (*thin lines* before adaptation, *thick lines* after adaptation; *triangle* indicates the adapting stimulus). *Right*: The *repulsive* tuning curve shifts correspond to a shift of the population activity (response of a population of neurons to a particular stimulus) *toward* the adapting stimulus. The neuron label is determined by each neuron's preferred orientation. (**b**) Same as (**a**), but now adaptation suppresses responses of individual neurons with preferred direction close to the adapting stimulus. *Right*: This corresponds to a shift of the population activity *away* from the adapting stimulus

observed at single neuron level in visual cortex is indeed consistent with the observed changes in perception (Jin et al. 2005). Note that in this framework it is commonly assumed that adaptation alters encoding in sensory cortical neurons, whereas downstream areas "reading out" the stimulus-related information are unaware of adaptation (Seriès et al. 2009). The predicted perceptual effect also depends on the hypothetical link between population activity and perception, that is, on the "read-out" strategy (such as peak activity, population vector, or maximum likelihood decoding).

More recently, there has been increasing interest in studying high-level adaptation effects, including face adaptation (Webster 2011) and adaptation to the perception of causal interactions (Rolfs et al. 2013). Computational modeling will be helpful in answering interesting questions related to these phenomena, such as to what degree high-level effects can be explained by adaptation to low-level features or how they arise across the processing hierarchy due to recurrent interactions between cortical areas.

Cross-References

- ▶ Perception, Bayesian Models of
- ▶ Perceptual Decision-Making
- ▶ Short-Term Plasticity, Biophysical Models
- ▶ Spike-Frequency Adaptation
- ▶ Visual Aftereffects, Models of

References

Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. Science 275: 220–224

Barlow HB (1961) Possible principles underlying the transformation of sensory messages. In: Rosenblith WA (ed) Sensor communication. MIT Press, Cambridge, MA, pp 217–234

Bednar JA, Miikkulainen R (2000) Tilt aftereffects in a self-organizing model of the primary visual cortex. Neural Comput 12:1721–1740

Benda J, Herz AVM (2003) A universal model for spike-frequency adaptation. Neural Comput 15:2523–2564. https://doi.org/10.1162/089976603322385063

Chelaru MI, Dragoi V (2008) Asymmetric synaptic depression in cortical networks. Cereb Cortex 18:771–788. https://doi.org/10.1093/cercor/bhm119

Clifford CW, Wenderoth P, Spehar B (2000) A functional angle on some after-effects in cortical vision. Proc Biol Sci 267:1705–1710. https://doi.org/10.1098/rspb.2000.1198

Clifford CW, Webster MA, Stanley GB et al (2007) Visual adaptation: neural, psychological and computational aspects. Vis Res 47:3125–3131. https://doi.org/10.1016/j.visres.2007.08.023

Compte A, Wang X-J (2006) Tuning curve shift by attention modulation in cortical neurons: a computational study of its mechanisms. Cereb Cortex 16:761–778. https://doi.org/10.1093/cercor/bhj021

Cortes JM, Marinazzo D, Series P et al (2011) The effect of neural adaptation on population coding accuracy. J Comput Neurosci. https://doi.org/10.1007/s10827-011-0358-4

Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. Nat Neurosci 8:1684–1689. https://doi.org/10.1038/nn1541

Doya K, Ishii S, Pouget A, Rao RPN (2007) Bayesian brain: probabilistic approaches to neural coding. MIT Press, Cambridge, MA

Dragoi V, Sharma J, Sur M (2000) Adaptation-induced plasticity of orientation tuning in adult visual cortex. Neuron 28:287–298

Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412:787–792. https://doi.org/10.1038/35090500

Jin DZ, Dragoi V, Sur M, Seung HS (2005) Tilt aftereffect and adaptation-induced changes in orientation tuning in visual cortex. J Neurophysiol 94:4038–4050. https://doi.org/10.1152/jn.00571.2004

King AJ, Dahmen JC, Keating P et al (2011) Neural circuits underlying adaptation and learning in the perception of auditory space. Neurosci Biobehav Rev 35: 2129–2139. https://doi.org/10.1016/j.neubiorev.2011.03.008

Kohn A (2007) Visual adaptation: physiology, mechanisms, and functional benefits. J Neurophysiol 97: 3155–3164. https://doi.org/10.1152/jn.00086.2007

Kvale MN, Schreiner CE (2004) Short-term adaptation of auditory receptive fields to dynamic stimuli. J Neurophysiol 91:604–612. https://doi.org/10.1152/jn.00484.2003

Laughlin S (1981) A simple coding procedure enhances a neuron's information capacity. Z Naturforsch C 36: 910–912

Maravall M, Petersen RS, Fairhall AL et al (2007) Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. PLoS Biol 5:e19. https://doi.org/10.1371/journal.pbio.0050019

Nagel KI, Doupe AJ (2006) Temporal processing and adaptation in the songbird auditory forebrain. Neuron 51:845–859. https://doi.org/10.1016/j.neuron.2006.08.030

Petersen RS, Panzeri S, Maravall M (2009) Neural coding and contextual influences in the whisker system. Biol Cybern 100:427–446. https://doi.org/10.1007/s00422-008-0290-5

Rolfs M, Dambacher M, Cavanagh P (2013) Visual adaptation of the perception of causality. Curr Biol CB 23:250–254. https://doi.org/10.1016/j.cub.2012.12.017

Sanchez-Vives MV, Nowak LG, McCormick DA (2000) Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. J Neurosci 20: 4267–4285

Seriès P, Stocker AA, Simoncelli EP (2009) Is the homunculus "aware" of sensory adaptation? Neural Comput 21:3271–3304. https://doi.org/10.1162/neco.2009.09-08-869

Stimberg M, Wimmer K, Martin R et al (2009) The operating regime of local computations in primary visual cortex. Cereb Cortex 19:2166–2180. https://doi.org/10.1093/cercor/bhn240

Stocker A, Simoncelli E (2006) Sensory adaptation within a Bayesian framework for perception. In: Weiss Y, Schölkopf B, Platt J (eds) Advances neural information process system, vol 18. MIT Press, Cambridge, MA, pp 1289–1296

Teich AF, Qian N (2003) Learning and adaptation in a recurrent model of V1 orientation selectivity. J Neurophysiol 89:2086–2100. https://doi.org/10.1152/jn.00970.2002

Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc Natl Acad Sci U S A 94: 719–723

Wark B, Lundstrom BN, Fairhall A (2007) Sensory adaptation. Curr Opin Neurobiol 17:423–429. https://doi.org/10.1016/j.conb.2007.07.001

Webster MA (2011) Adaptation and visual coding. J Vis 11:1–23. https://doi.org/10.1167/11.5.3

Adaptive Control

▶ Spinal Cord, Integrated (Non CPG) Models of

Adaptive Design Optimization

▶ Adaptive Stimulus Optimization

Adaptive Sampling

▶ Adaptive Stimulus Optimization

Adaptive Stimulus Optimization

Christopher DiMattina[1] and Kechen Zhang[2]
[1]Department of Psychology, Florida Gulf Coast Univerisity, Fort Myers, FL, USA
[2]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

Synonyms

Adaptive design optimization; Adaptive sampling; Closed-loop experiments; Optimal experimental design; Optimal stimulus design

Definition

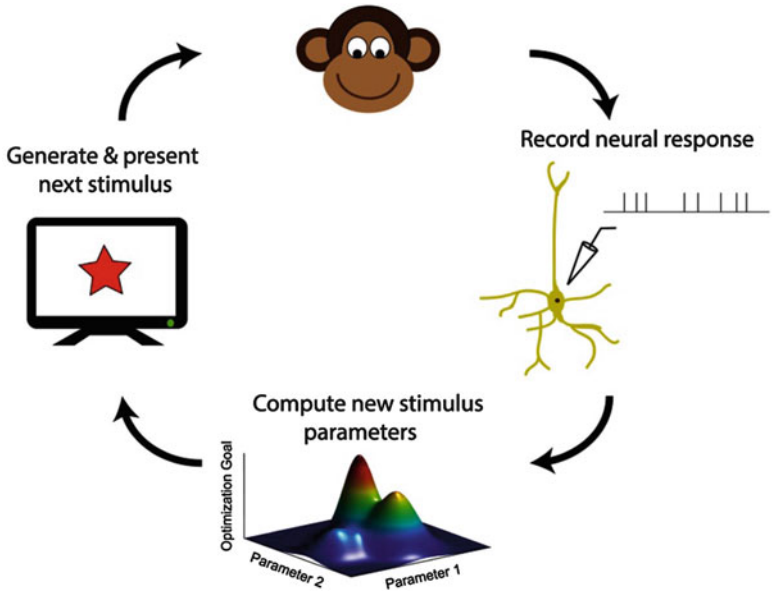
Adaptive stimulus optimization refers to an experimental approach in neuroscience where neuronal or behavioral responses to stimuli presented on previous trials are utilized to adaptively generate new stimuli in an iterative, closed-loop manner, usually by optimizing an objective function. There are different choices for the objective function. For example, if the objective function is the neural response itself, the optimization procedure finds an optimal stimulus that drives maximum response or is at least a local optimum in the stimulus space. When the objective function is the mutual information between the responses and the unknown parameters of a stimulus-response model, the optimization finds the stimulus set that yields the most accurate parameter estimation.

Detailed Description

Overview

Traditional experiments in the neurosciences have typically a fixed set of stimuli chosen *a priori* to elicit responses from neurons in an open-loop paradigm, with data analysis and model fitting taking place post hoc. In recent years, with increases in computer power and improvements of algorithms, there has been a growing interest in adaptively generating stimuli online during the course of experimentation in an iterative, closed-loop manner, where neuronal responses from previous trials are used to generate new stimuli (Benda et al. 2007; DiMattina and Zhang 2013; Potter et al. 2013; Park and Pillow 2016). This general paradigm is illustrated schematically in Fig. 1.

Adaptive stimulus optimization has long been used in psychophysics for estimating sensory thresholds (Watson and Pelli 1983; Kontsevich and Tyler 1999) and enjoys a large body of theoretical results from the statistics and machine learning literature (Paninski 2005; Chaloner and Verdinelli 1995). In sensory neuroscience studies, stimuli have been adaptively optimized for a wide variety of experimental goals, including maximizing neural firing rates (O'Connor et al. 2005; Chambers et al. 2014), finding maximally informative stimulus ensembles, (Machens et al. 2005), and estimating and comparing models of sensory processing (Lewi et al. 2009; DiMattina and Zhang 2011; Tam 2012; Park and Pillow 2012, 2016). In addition to applications in systems-level sensory neuroscience, closed-loop approaches have also been applied in many diverse areas including cognitive science, cellular neurophysiology, and brain-computer interfaces (Myung et al. 2013; Potter et al. 2013).



Adaptive Stimulus Optimization, Fig. 1 Schematic illustration of adaptive stimulus optimization, where responses to preceding stimuli are used to generate subsequent stimuli in a closed-loop manner (Reproduced from DiMattina and Zhang 2013)

Optimizing Firing Rate

Methods for adaptive optimization of neuronal firing rate fall broadly into two categories: (1) hill-climbing methods and (2) genetic algorithms. Hill-climbing methods utilize local perturbations of a reference stimulus to estimate the local response surface from noisy neural responses, iteratively moving the reference stimulus in a direction (e.g., the gradient) which increases neural firing rate (Harth and Tzanakou 1974; O'Connor et al. 2005; Nelken et al. 1994; Koelling and Nykamp 2012). Genetic algorithms mimic biological evolution by broadly populating the stimulus space with numerous stimuli and using their elicited neural responses as a measure of fitness. The fittest stimuli in each generation are then used to define the next generation of stimuli by recombination of their features in a manner analogous to sexual reproduction (Yamane et al. 2008; Chambers et al. 2014). Genetic algorithms have the advantage of being more robust to local maxima than hill-climbing methods and more extensively sampling the stimulus space.

Iso-response Surfaces

Instead of finding the single stimulus that optimizes the firing rate, it is also useful to find the set of all stimuli which elicit the same firing rate response. The shape of these firing rate level sets can tell us about how a sensory neuron combines stimulus dimensions. This method has been applied in diverse contexts, including studies of spectral integration in grasshopper auditory neurons and integration of photoreceptor inputs by V1 neurons (Gollisch et al. 2002; Horwitz and Hass 2012).

Optimizing Information

Instead of characterizing a neuron by its preferred or "optimal" stimulus, an alternative approach is to characterize the neuron in terms of the stimulus ensemble which its responses most reliably distinguish. This may be quantified by maximizing the mutual information between the stimuli and neural responses, and this technique was applied by Machens et al. (2005) in a study of grasshopper auditory receptor neurons.

Estimating and Comparing Models

Given an accurate model of the input-output relationship for a sensory neuron, it is possible in principle to predict the neuron's response to an arbitrary stimulus. However, estimating high-dimensional models from limited experimental data often poses a serious technical challenge. A study by Lewi et al. (2009) demonstrated that adaptively selecting stimuli to optimize expected

mutual information between neural responses and model parameters allowed fast and robust estimation of generalized linear models. Subsequent work by DiMattina and Zhang (2011) extended this idea to arbitrary stimulus-response models and also considered the problem of adaptively optimizing stimuli for comparing multiple, competing neural models. These methods were verified experimentally in a study of spectral integration in the primate inferior colliculus (Tam 2012). More recent work has considered the use of well-chosen priors to further speed convergence of receptive field estimates (Park and Pillow 2012, 2016). Optimization of sensory stimuli for model estimation and comparison have also been recently applied in vision psychophysics and cognitive science (Wang and Simoncelli 2008; Myung et al. 2013; Kim et al. 2014).

Cross-References

- Bayesian Approaches in Computational Neuroscience: Overview
- Estimation of Neuronal Firing Rate
- Information Theory: Overview
- Neural Coding
- Spectrotemporal Receptive Fields

References

Benda J, Gollisch T, Machens CK, Herz AV (2007) From response to stimulus: adaptive sampling in sensory physiology. Curr Opin Neurobiol 17(4):430–436

Chaloner K, Verdinelli I (1995) Bayesian experimental design: a review. Stat Sci 10(3):273–304

Chambers AR, Hancock KE, Sen K, Polley DB (2014) Online stimulus optimization rapidly reveals multidimensional selectivity in auditory cortical neurons. J Neurosci 34(27):8963–8975

DiMattina C, Zhang K (2011) Active data collection for efficient estimation and comparison of nonlinear neural models. Neural Comput 23(9):2242–2288

DiMattina C, Zhang K (2013) Adaptive stimulus optimization for sensory systems neuroscience. Front Neural Circuits 7:101

Gollisch T, Schu¨tze H, Benda J, Herz AV (2002) Energy integration describes sound-intensity coding in an insect auditory system. J Neurosci 22(23):10434–10448

Harth E, Tzanakou E (1974) Alopex: a stochastic method for determining visual receptive fields. Vis Res 14(12):1475–1482

Horwitz GD, Hass CA (2012) Nonlinear analysis of macaque v1 color tuning reveals cardinal directions for cortical color processing. Nat Neurosci 15(6):913–919

Kim W, Pitt MA, Lu Z-L, Steyvers M, Myung JI (2014) A hierarchical adaptive approach to optimal experimental design. Neural Comput 26(11):2465–2492

Koelling ME, Nykamp DQ (2012) Searching for optimal stimuli: ascending a neurons response function. J Comput Neurosci 33(3):449–473

Kontsevich LL, Tyler CW (1999) Bayesian adaptive estimation of psychometric slope and threshold. Vis Res 39(16):2729–2737

Lewi J, Butera R, Paninski L (2009) Sequential optimal design of neurophysiology experiments. Neural Comput 21(3):619–687

Machens CK, Gollisch T, Kolesnikova O, Herz AV (2005) Testing the efficiency of sensory coding with optimal stimulus ensembles. Neuron 47(3):447–456

Myung JI, Cavagnaro DR, Pitt MA (2013) A tutorial on adaptive design optimization. J Math Psychol 57(3):53–67

Nelken I, Prut Y, Vaadia E, Abeles M (1994) In search of the best stimulus: an optimization procedure for finding efficient stimuli in the cat auditory cortex. Hear Res 72(1):237–253

O'Connor KN, Petkov CI, Sutter ML (2005) Adaptive stimulus optimization for auditory cortical neurons. J Neurophysiol 94(6):4051–4067

Paninski L (2005) Asymptotic theory of information-theoretic experimental design. Neural Comput 17(7):1480–1507

Park M, Pillow JW (2012) Bayesian active learning with localized priors for fast receptive field characterization. In: Advances in neural information processing systems. Cambridge, MA: MIT Press, 25:2348–2356

Park M, Pillow JW (2016) Adaptive bayesian methods for closed-loop neurophysiology. In: El Hady A (ed) Closed loop neuroscience. Cambridge, MA: Academic Press/Elsevier, pp 3–16

Potter SM, El Hady A, Fetz EE (2013) Closed-loop neuroscience and neuroengineering. Front Neural Circuits 8:115

Tam W (2012) Adaptive modeling of marmoset inferior colliculus neurons in vivo. PhD thesis, The Johns Hopkins University School of Medicine

Wang Z, Simoncelli EP (2008) Maximum differentiation (mad) competition: a methodology for comparing computational models of perceptual quantities. J Vis 8(12):8

Watson AB, Pelli DG (1983) Quest: a bayesian adaptive psychometric method. Percept Psychophys 33(2):113–120

Yamane Y, Carlson ET, Bowman KC, Wang Z, Connor CE (2008) A neural code for three-dimensional object shape in macaque inferotemporal cortex. Nat Neurosci 11(11):1352–1360

Afferent Feedback to Neural Oscillators

► Sensory Input to Central Pattern Generators

Afferent Input to Rhythm Generating Networks

► Sensory Input to Central Pattern Generators

Algorithmic Generation of Motoneuron Morphology

► Algorithmic Reconstruction of Motoneuron Morphology

Algorithmic Reconstruction of Motoneuron Morphology

Joseph Graham
Blue Brain Project, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Synonyms

Algorithmic generation of motoneuron morphology; Computational synthesis of motoneuron morphology; Computer generation of motoneuron morphology

Definition

Algorithmic reconstruction of neuronal morphology is the process of parameterizing neurite branching patterns, quantifying the parameters for a given population of experimentally reconstructed neurons, and then feeding the data into an algorithm which computationally generates populations of "virtual" neurons.

Detailed Description

Background

Digitization of neuronal morphology is important for computational neuroscience because neuronal morphology affects synaptic integration and firing behavior within individual neurons as well as determining potential connectivity with other neurons (Ascoli 2002). However, experimental reconstruction techniques are still largely manual or, at best, semiautomated, requiring time and skill to accurately capture neuronal morphology. As computational models of the nervous system grow in scale with computational power, there is an increasing need for large numbers of digitized ("virtual") morphologies. Algorithmic generation of virtual morphologies has the potential to fulfill this need.

Motoneurons, as the nervous system's "final common pathway" of motor control (Sherrington 1906), have long been studied. These neurons exhibit extensive dendritic arborizations that may stretch over several millimeters in the spinal cord (Cullheim et al. 1987a, b) which makes their reconstruction particularly challenging. As such, the earliest forays into algorithmic generation of neuronal morphology occurred in motoneurons. In particular, one set of reconstructed motoneurons from Cullheim et al. (1987a, b) has been the basis of algorithmic generation research by multiple groups (these morphologies are available for download at neuromorpho.org).

The general process of algorithmic generation begins by parameterizing the neurite branching patterns. These parameters are statistical distributions which are then quantified from experimental reconstructions. Generation then proceeds from the soma outwards, first generating primary neurites which go on to branch or terminate according to the chosen algorithm. This process continues until all branches have terminated. The algorithmic reconstructions are then compared to the experimental reconstructions, whereupon persistent differences are explored to improve the parameterization and algorithm. As this process is repeated, the algorithmic reconstructions become more and more similar to the experimental. The ultimate goal of this line of research is the

capability of generating populations of unique morphologies which are statistically indistinguishable from the experimental population upon which they are based and which capture the natural variability of such a population.

Parameterization

Hillman (1979) was the first to propose that a set of "fundamental" parameters could be used to completely and parsimoniously describe neuronal morphology. He recognized that there are two separable aspects of neuronal morphology that must be parameterized: the branching patterns and the space-filling behavior. It is important to remember that these parameters are not scalar values, but rather statistical distributions, and also that these parameters may be intercorrelated or correlated with other local properties of the arborization.

To capture the branching patterns, Hillman proposed that every neurite tree begins with a "stem diameter" and grows for a certain "segment length" while changing diameter by "segment taper." If the final diameter of a branch is larger than the "terminal diameter," the branch bifurcates into two daughter branches, whose diameters are related to that of their parent by the "branch power" and whose relative diameters are determined by the "daughter ratio." Hillman proposed that quantifying these parameters was sufficient to completely describe neurite branching patterns.

To capture the space-filling patterns, Hillman proposed that one must measure the initial direction of neurite trees and the branching angles (the angles that daughter branches make with regard to the parent branch). Since that time, it has been recognized that to capture space filling, more parameters need to be measured. Proposed additions include some measure of "meander" (how much a neurite's direction changes as a branch extends outwards) and some measure of "tropism" (a force inducing a directionality in neurite behavior, such as an apical dendrite extending towards the pia).

Slightly different parameter sets have since been proposed and explored based on Hillman's original fundamental parameters.

Algorithms

Burke et al. (1992) were the first to realize that parameterization of neuronal morphology could be used to algorithmically generate virtual neurite trees. Using experimentally reconstructed motoneurons (Cullheim et al. 1987a, b), their strategy was to "(1) devise a model system that can simulate dendritic trees, (2) derive the required model parameters directly from measurements of real dendrites, and (3) refine the parameter derivations or basic model assumptions, based on the degree of congruence between real and simulated dendrites" (Burke et al. 1992). Using this methodology, Burke et al. found correlations between their parameters and local diameter and path length from the soma. Marks and Burke later extended this work (2007a, b) to include space-filling parameters and the generation of entire neurons.

Ascoli and Krichmar (2000) were the first to apply the methodology in order to create entire neurons, by adding somatic and neurite trunk parameters to a software package called L-Neuron. This software package could replicate Burke et al.'s (1992) algorithm, as well as several algorithms derived more closely from Hillman's (1979) parameterization. Ascoli et al. (2001) then went on to explore these algorithms using experimental reconstructions of several different types of neurons, including the motoneurons from Cullheim et al. (1987a, b). Donohue and Ascoli (2008) later extended this work to allow parameters to be correlated with local neurite properties (branch order, diameter, and path length from the soma). In an extensive analysis, they explored which local properties are most important for generating realistic morphologies.

Many groups have been exploring algorithmic generation of neuronal morphology, but relatively few have explicitly explored motoneurons. Torben-Nielsen et al. (2008) developed a very different sort of algorithm which they used to explore motoneuron morphology. In this algorithm, parameters are not quantified in advance, but rather the entire parameter space is explored in an evolutionary algorithm which can converge on parameter values which produce realistic morphologies.

Summary

Algorithmic generation of motoneuron morphology is a subset of general algorithmic generation of neuronal morphology. The methodology offers the promise of a deeper understanding of neuronal morphology and may eventually fulfill the need of computational neuroscientists for large numbers of realistic morphologies for use in simulations. While no algorithm has yet been capable of producing morphologies which are statistically indistinguishable from experimental reconstructions across all morphometrics, much progress has been made. The study of motoneurons is especially useful for algorithm development because several different groups have explored the same data set (Cullheim et al. 1987a, b), thus providing a basis for further improvement.

Cross-References

- ▶ Compartmental Models of Spinal Motoneurons
- ▶ Morphologically Detailed Cellular and Pool Motoneuron Models
- ▶ Reconstruction, Techniques and Validation
- ▶ Synthetic Neuronal Circuits/Networks
- ▶ Synthetic Neuronal Morphology

References

Ascoli GA (ed) (2002) Computational neuroanatomy: principles and methods. Humana Press, Totowa

Ascoli GA, Krichmar JL (2000) L-neuron: a modeling tool for the efficient generation and parsimonious description of dendritic morphology. Neurocomputing 32–33: 1003–1011

Ascoli GA, Krichmar JL, Scorcioni R, Nasuto SJ, Senft SL (2001) Computer generation and quantitative morphometric analysis of virtual neurons. Anat Embryol 204: 283–301

Burke RE, Marks WB, Ulfhake B (1992) A parsimonious description of motoneuron dendritic morphology using computer simulation. J Neurosci 12(6):2403–2416

Cullheim S, Fleshman JW, Glenn LL, Burke RE (1987a) Membrane area and dendritic structure in type-identified triceps surae alpha motoneurons. J Comp Neurol 255:68–81

Cullheim S, Fleshman JW, Glenn LL, Burke RE (1987b) Three-dimensional architecture of dendritic trees in type-identified alpha motoneurons. J Comp Neurol 255:82–96

Donohue DE, Ascoli GA (2008) A comparative computer simulation of dendritic morphology. PLoS Comput Biol 4(5):e1000089

Hillman DE (1979) Neuronal shape parameters and substructures as a basis of neuronal form. In: Schmitt FO, Worden FG (eds) The neurosciences: fourth study program. MIT Press, Cambridge, MA, pp 477–498

Marks WB, Burke RE (2007a) Simulation of motoneuron morphology in three dimensions. I Building individual dendritic trees. J Comp Neurol 503:685–700

Marks WB, Burke RE (2007b) Simulation of motoneuron morphology in three dimensions. II Building complete neurons. J Comp Neurol 503:701–716

Sherrington CS (1906) Integrative actions of the nervous system. Yale University Press, New Haven

Torben-Nielsen B, Tuyls K, Postma E (2008) EvOL-Neuron: neuronal morphology generation. Neurocomputing 71: 963–972

Algorithms for Generating Neuronal Morphologies

- ▶ Synthetic Neuronal Morphology

Amari Model

Roland Potthast
Department of Mathematics, University of Reading, Reading, UK

Definition

The Amari neural field model (cf. (Amari 1975, 1977)) provides a simple field-theoretic approach to the dynamics of neural activity in the brain. The model uses excitations and inhibitions over some distance as an effective model of mixed inhibitory and excitatory neurons with typical cortical connectivities. The model is a scalar dynamical equation for the voltage or activity $u(x, t)$ of the form.

$$\frac{\partial u}{\partial t}(x,t) = -u(x,t) + \int_B w(x,y) f(u(y,t)) dy, x \in B, t \geq 0, \quad (1)$$

where initial conditions $u(x, 0) = u_0(x), x \in B$ are given. Here, B is our *brain*, i.e., some domain where the neural activity takes place; f is the local *activation function* or *firing rate function*; and w is the *connectivity function* which models the strength of the connectivity or signal propagation from $y \in B$ to the point x.

A common choice for the activation function has *sigmoidal shape.*

$$f(s) := \frac{1}{1 + e^{-\beta(s-h)}}, s \in \mathbb{R}, \tag{2}$$

which is monotonously growing from $f(-\infty) = 0$ to saturation $f(\infty) = 1$ for large s, where h is the *threshold* and β is a *steepness parameter*. Often, the kernel w is chosen to be homogeneous, i.e., $w(x, y) = w(x - y)$. In this case, the integral becomes a convolution integral, such that the Amari equation can be written in the form.

$$\frac{\partial u}{\partial t} = -u + w \otimes f(u). \tag{3}$$

The Amari equation provides a continuous analogue or continuous description of *neural networks*, which has become widely used in the engineering community.

Historic Background and Applications

The earliest *field models* for describing and studying neural activity dynamics go back to Beurle (1956), investigating the proportion of active neurons in randomly connected networks, followed by work of Griffith (1963, 1965). The basis of modern field dynamical models has been the work of Cowan, Nunez, and Amari in the 1970s (see Wilson and Cowan 1972, 1973; Nunez 1974; Amari 1977). Cowan proposed an activity-based model with two distinct populations of excitatory and inhibitory subpopulations. Amari suggested a more condensed scalar model with a Mexican hat-type connectivity function, where excitation and inhibition are reflected by the changing sign of the connectivity kernel $w(x - y)$.

The Amari neural field model has been applied, e.g., to autonomous robotic behavior (Erlhagen and Bicho 2006), embodied cognition (Schöner and Dineva 2007), dynamic causal modeling (Daunizeau et al. 2009), and language processing (beim Graben et al. 2008; beim Graben and Potthast 2012); for further details, we refer to the recent tutorial by Coombes et al. (2013).

Theory

The problem (4) is an integrodifferential equation, establishing an *evolutionary dynamical system* with initial condition $u(x, 0) = u_0(x), x \in B$, at time $t = 0$. For a Lipschitz continuous activation function f, *existence* and *uniqueness* of a solution u of Eq. 4 for all times and in one or several dimensions are obtained under quite general conditions based on elementary arguments and the fixed-point theorem (compare Potthast and beim Graben 2010). Over time, there has been significant activity to study the existence and uniqueness of *bumps* and *waves* in one spatial dimension when a particular homogeneous kernel $w(x - y)$ is given (see Kishimoto and Amari 1979; Ermentrout and McLeod 1993) where smooth sigmoidal firing rates are used (see also Coombes and Schmidt 2010; Oleynik et al. 2013; Coombes and Owen 2004 for more general but still rather restricted classes of functions f).

Delay Neural Field Equation and Homogeneous Kernels

Geometric singular perturbation analysis investigates the stability of static or transient solutions, and *numerical bifurcation techniques* investigate the existence and properties of bifurcation points in state or parameter spaces (compare Pinto and Ermentrout 2001a, b; Laing and Troy 2003b). For studies in two or more dimensions, we refer to Taylor (1999), Laing and Troy (2003a), Folias and Bressloff (2004), Laing (2005), Owen et al. (2007), Kilpatrick and Bressloff (2010a), and

Coombes et al. (2012). Often, the Eq. 4 is complemented by a *delay term.*

$$\frac{\partial u}{\partial t}(x,t) = -u(x,t) + \int_B w(x,y) f\left(u\left(y, y - \frac{D(x,y)}{v}\right)\right) dy, \tag{4}$$

$x \in B, t \geq 0$, where $D(x, y)$ is the length of the fiber between x and y and v is the finite propagation speed of signals. The above analysis for neural field equations has been extended to delay equations (see Nunez 1974; Jirsa and Haken 1997; Coombes et al. 2003; Hutt 2004; Venkov et al. 2007; Grindrod and Pinotsis 2010), including dendritic processing (Bressloff and Coombes 1997) and synaptic depression (Kilpatrick and Bressloff 2010b).

Most of the above work uses homogeneous kernels $w(x - y)$. More recent work that tackles heterogeneity (primarily using simulations) can be found in Brackley and Turner (2007), Bressloff (2012), Schmidt et al. (2009), and Coombes et al. (2012) and functional analytic results in Faugeras et al. (2008), and Potthast and beim Graben (2010). The *inverse problems* perspective for either homogeneous or nonhomogeneous kernels has been investigated by Potthast and beim Graben (2009) and beim Graben and Potthast (2009). More recently, *stochastic neural field equations* have become very popular; compare the review by Bressloff (2012).

Advantages of the Approach

The Amari neural field model provides a very concise scalar equation to model neural activity and dynamics. In contrast to microscopic models, it summarizes the activity of neurons into the activity function $u(x, t)$, which can be used to reduce the computational complexity significantly. The field-theoretic approach opens the dynamic system to mathematical analysis and beyond the range of discrete network models.

Limitations

The strength of the Amari model is its simplicity, which is at the same time its strongest limitation. It does not take into account the complex chemical and physiological processes which take place in addition to electrical dynamics in neural tissue, nor is it capable to include the different temporal scales on which these processes work and propagate.

Cross-References

- Bifurcations, Neural Population Models and
- Chaos, Neural Population Models and
- Inverse Problems in Neural Population Models
- Neural Field Model, Continuum
- Neural Population Model
- Pattern Formation in Neural Population Models
- Phase Transitions, Neural Population Models and
- Stochastic Neural Field Theory
- Wilson-Cowan Model

References

Amari S (1975) Homogeneous nets of neuron-like elements. Biol Cybern 17:211–220

Amari S (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybern 27:77–87

Beim Graben P, Potthast R (2009) Inverse problems in dynamic cognitive modeling. Chaos 19(1):015103

Beim Graben P, Potthast R (2012) A dynamic field account to language-related brain potentials. In: Rabinovich M, Friston K, Varona P (eds) Principles of brain dynamics: global state interactions. MIT Press, Cambridge, MA

Beim Graben P, Pinotsis D, Saddy D, Potthast R (2008) Language processing with dynamic fields. Cogn Neurodyn 2(2):79–88

Beurle RL (1956) Properties of a mass of cells capable of regenerating pulses. Phil Trans R Soc Lond B 240: 55–94

Brackley CA, Turner MS (2007) Random fluctuations of the firing rate function in a continuum neural field model. Phys Rev E 75:041913

Bressloff PC (2012) Spatiotemporal dynamics of continuum neural fields. J Phys A 45:033,001

Bressloff PC, Coombes S (1997) Physics of the extended neuron. Int J Mod Phys B 11:2343–2392

A

Coombes S, Owen MR (2004) Evans functions for integral neural field equations with Heaviside firing rate function. SIAM J Appl Dyn Syst 34:574–600
Coombes S, Schmidt H (2010) Neural fields with sigmoidal firing rates: approximate solutions. Discret Contin Dyn Syst Ser A 28:1369–1379
Coombes S, Lord GJ, Owen MR (2003) Waves and bumps in neuronal networks with axo-dendritic synaptic interactions. Phys D 178:219–241
Coombes S, Schmidt H, Bojak I (2012) Interface dynamics in planar neural field models. J Math Neurosci 2:9
Coombes S, Beim Graben P, Potthast R et al (2013) Tutorial on neural field theory. In: Wright J, Potthast R, Coombes S, Beim Graben P (eds) Neural fields. Theory and applications. Springer, Berlin
Daunizeau J, Kiebel SJ, Friston KJ (2009) Dynamic causal modelling of distributed electromagnetic responses. NeuroImage 47:590–601
Erlhagen W, Bicho E (2006) The dynamic neural field approach to cognitive robotics. J Neural Eng 3:R36–R54
Ermentrout GB, McLeod JB (1993) Existence and uniqueness of travelling waves for a neural network. Proc Roy Soc Edinb 123A:461–478
Faugeras O, Grimbert F, Slotine JJ (2008) Absolute stability and complete synchronization in a class of neural fields models. SIAM J Appl Math 69:205–250
Folias SE, Bressloff PC (2004) Breathing pulses in an excitatory neural network. SIAM J Appl Dyn Syst 3:378–407
Griffith JS (1963) A field theory of neural nets: I: derivation of field equations. Bull Math Biophys 25:111–120
Griffith JS (1965) A field theory of neural nets: II: properties of field equations. Bull Math Biophys 27:187–195
Grindrod P, Pinotsis D (2010) On the spectra of certain integro-differential-delay problems with applications in neurodynamics. Phys D Nonlinear Phenom 240(1):13–20. https://doi.org/10.1016/j.physd.2010.08.002. ISSN 0167–2789
Hutt A (2004) Effects of nonlocal feedback on traveling fronts in neural fields subject to transmission delay. Phys Rev E 60(1–4):052902
Jirsa VK, Haken H (1997) A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. Phys D 99:503–526
Kilpatrick ZP, Bressloff PC (2010a) Effects of synaptic depression and adaptation on spatiotemporal dynamics of an excitatory neuronal network. Phys D 239:547–560
Kilpatrick ZP, Bressloff PC (2010b) Spatially structured oscillations in a two-dimensional excitatory neuronal network with synaptic depression. J Comput Neurosci 28:193–209
Kishimoto K, Amari S (1979) Existence and stability of local excitations in homogeneous neural fields. J Math Biol 7:303–318
Laing CR (2005) Spiral waves in nonlocal equations. SIAM J Appl Dyn Syst 4:588–606
Laing CR, Troy WC (2003a) PDE methods for nonlocal models. SIAM J Appl Dyn Syst 2:487–516
Laing CR, Troy WC (2003b) Two bump solutions of Amari-type models of working memory. Phys D 178:190–218
Nunez PL (1974) The brain wave equation: a model for the EEG. Math Biosci 21:279–297
Oleynik A, Posnov A, Wyller J (2013) On the properties of nonlinear nonlocal operators arising in neural field models. J Math Anal Appl 398:335–351
Owen MR, Laing CR, Coombes S (2007) Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. New J Phys 9:378
Pinto DJ, Ermentrout GB (2001a) Spatially structured activity in synaptically coupled neuronal networks: I. Travelling fronts and pulses. SIAM J Appl Math 62:206–225
Pinto DJ, Ermentrout GB (2001b) Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses. SIAM J Appl Math 62:226–243
Potthast R, Beim Graben P (2009) Inverse problems in neural field theory. SIAM J Appl Dyn Syst 8(4):1405–1433
Potthast R, Beim Graben P (2010) Existence and properties of solutions for neural field equations. Math Methods Appl Sci 33(8):935–949
Schmidt H, Hutt A, Schimansky-Geier L (2009) Wave fronts in inhomogeneous neural field models. Phys D 238:1101–1112
Schöner G, Dineva E (2007) Dynamic instabilities as mechanisms for emergence. Dev Sci 10:69–74
Taylor JG (1999) Neural 'bubble' dynamics in two dimensions: foundations. Biol Cybern 80:393–409. Structure
Venkov NA, Coombes S, Matthews PC (2007) Dynamic instabilities in scalar neural field equations with space-dependent delays. Phys D 232:1–15
Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J 12:1–24
Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13:55–80

Further Reading

Ben-Yishai R, Bar-Or L, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. Proc Natl Acad Sci U S A 92:3844–3848
Berger H (1929) Über das Elektroenkephalogramm des Menschen. Archiv Psychiatr 87:527–570
Bressloff PC (2001) Traveling fronts and wave propagation failure in an in-homogeneous neural network. Phys D 155:83–100
Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener M (2001) Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. Phil Trans R Soc Lond B 40:299–330
Ermentrout GB, Cowan JD (1979) A mathematical theory of visual hallucination patterns. Biol Cybern 34:137–150
Geise MA (1999) Neural field theory for motion perception. Kluwer, Boston

Jirsa VKV, Jantzen KJ, Fuchs A, Kelso JAS (2001) Information processing in medical imaging, chap. Neural field dynamics on the folded three-dimensional cortical sheet and its forward EEG and MEG. Springer, Berlin, pp. 286–299
Jirsa VK, Jantzen KJ, Fuchs A, Kelso JAS (2002) Spatiotemporal forward solution of the EEG and MEG using network modeling. IEEE Trans Med Imaging 21(5): 493–504
Kilpatrick ZP, Bressloff PC (2010) Binocular rivalry in a competitive neural network with synaptic depression. SIAM J Appl Dyn Syst 9:1303–1347
Laing CR, Troy WC, Gutkin B, Ermentrout GB (2002) Multiple bumps in a neuronal model of working memory. SIAM J Appl Math 63:62–97
Liley DTJ, Cadusch PJ, Dafilis MP (2002) A spatially continuous mean field theory of electrocortical activity. Network 13:67–113
Liley DTJ, Foster BL, Bojak I (2011) Sleep and anesthesia, chap. A mesoscopic modelling approach to anaesthetic action on brain electrical activity. Springer, New York, pp. 139–166
Nunez PL (1995) Neocortical dynamics and human EEG rhythms. Oxford University Press, New York
Rabinovich M, Friston K, Varona P (eds) (2012) Principles of brain dynamics: global state interactions. MIT Press, Cambridge, MA
Tass P (1995) Cortical pattern formation during visual hallucinations. J Biol Phys 21:177–210

Ambiguous Stimuli

▶ Multistability in Perception Dynamics

AMPA Glutamate Receptor (AMPA Receptor), Conductance Models

Patrick D. Roberts
Department of Biomedical Engineering,
Oregon Health and Science University, Portland, OR, USA

Definition

A glutamate receptor that is permeable to sodium ions and carries an excitatory synaptic current following the binding of glutamate. AMPA receptors are regulated to control the maximum synaptic current by processes of synaptic plasticity and they co-localize with NMDA receptors.

Detailed Description

AMPA receptors are named for the selective agonist (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) that does not bind well to other glutamate receptors. The receptor is permeable to cations and can allow Na^+, K^+, and Ca^{2+} to cross the membrane and has an equilibrium potential $E_{AMPA} = 0\ mV$. The permeability to Ca^{2+} is small and is not considered important for initiating signaling cascades. AMPA receptors are composed of four types of subunits, and the combination of subunits determines the kinetics and permeability to cations.

Kinetics of AMPA Receptors

Mathematical representations of AMPA currents include the time-dependent synaptic conductance, $g_{AMPA}(t)$, in the current–voltage equation:

$$I_{AMPA}(t) = g_{AMPA}(t)(V(t) - E_{AMPA}) \quad (1)$$

The time-dependent synaptic conductance represents the opening and closing kinetics of the receptor channels and may be based on a double exponential function (or a similarly shaped function).

$$g_s(t) = \bar{g}_s\left(e^{-t/\tau_1} - e^{-t/\tau_2}\right) \quad (2)$$

where τ_1 is the onset time constant ($\tau_2 < 0.1$ ms) and τ_2 is the decay time constant ($\tau_2 < 3$ ms) (Hestrin et al. 1990; Trussell et al. 1993; Jonas et al. 1993). AMPA currents have a faster decay than NMDA currents (Jonas and Spruston 1994) and typically have a substantially larger peak current.

Because AMPA currents are usually the dominant excitatory synaptic currents, simple representations of these currents may be the only synaptic currents included in the practice of

modeling neural circuits. In models with low time resolution, AMPA currents may be represented as increased currents in a single time bin without short-term temporal details. However, these simple models may contain long-term plasticity that changes the circuit dynamics over time by changing the strength of connections between neurons.

Computational Functions of AMPA Receptors

The excitatory characteristics of AMPA currents can be the main driver of activity in a network of neurons and for communication across long distances such as from the periphery to the central nervous system. Due to the short time constants of AMPA kinetics, signal transmission can carry high temporal precision. An example of this precision is found in the auditory pathway where the interaural time difference can be discerned to extreme precision (Konishi 1990).

Another consequence of the short duration of AMPA currents is that they require a large population of independent asynchronous inputs to deliver a constant depolarization to a neuron. A single AMPA synapse can deliver a fast synaptic current that is filtered by cable properties of the dendrite where the postsynaptic terminal is located (Rall 1967). But there are temporal limitations to how much filtering is possible to deliver the excitatory effect to the spike-generating zone of the postsynaptic neuron. Multiple AMPA currents located on the postsynaptic neurons can overcome this constraint if the rate of incoming spikes is high enough and well distributed in time to overlap. Thus, the synaptic dynamics of a single synaptic terminal may not have a great influence on the postsynaptic activity.

The strength of the synaptic current is not the only determinant of the efficacy of the synaptic input to generate a spike in the postsynaptic neuron. Spikes are triggered by changes in membrane potential, and the filtering properties of the neuron are critical in how the dynamics of the individual synapses are transformed into changes in postsynaptic activity.

AMPA Synaptic Dynamics

In addition to the channel kinetics, AMPA currents exhibit short-term plasticity that can result from both pre- and postsynaptic mechanisms (Zucker and Regehr 2002; Blitz et al. 2004). Presynaptic mechanisms affect the release of glutamate and can influence the peak of the current on subsequent presynaptic spikes. Postsynaptic mechanisms affect the response of AMPA receptors to the concentration of glutamate in the synaptic cleft.

AMPA currents are also involved in long-term plasticity and can be the main component of changes in synaptic strength. AMPA currents do not play a direct role in inducing long-term plasticity but reflect the changes in presynaptic release and their own response to glutamate caused by other mechanisms.

Cross-References

- ▶ Kinetic Models of Postsynaptic Currents
- ▶ N-Methyl-D-Aspartate (NMDA) Receptors, Conductance Models

References

Blitz DM, Foster KA, Regehr WG (2004) Short-term synaptic plasticity: a comparison of two synapses. Nat Rev Neurosci 5(8):630–640

Hestrin S, Sah P, Nicoll RA (1990) Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. Neuron 5(3): 247–253

Jonas P, Spruston N (1994) Mechanisms shaping glutamate-mediated excitatory postsynaptic currents in the CNS. Curr Opin Neurobiol 4(3):366–372

Jonas P, Major G, Sakmann B (1993) Quantal components of unitary EPSCs at the mossy fibre synapse on CA3 pyramidal cells of rat hippocampus. J Physiol 472(1): 615–663

Konishi M (1990) The neural algorithm for sound localization in the owl. Harvey Lect 86:47

Rall W (1967) Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input. J Neurophysiol 30(5):1138

Trussell LO, Zhang S, Ramant IM (1993) Desensitization of AMPA receptors upon multiquantal neurotransmitter release. Neuron 10(6):1185–1196

Zucker RS, Regehr WG (2002) Short-term synaptic plasticity. Annu Rev Physiol 64(1):355–405

Further Reading

Dayan P, Abbott LF, Abbott L (2001) Theoretical neuroscience: computational and mathematical modeling of neural systems. Taylor & Francis, Cambridge, MA

Koch C (2004) Biophysics of computation: information processing in single neurons. Oxford university press, New York

Amplitude-Amplitude Coupling

▶ Theta-Gamma Cross-Frequency Analyses (Hippocampus)

Anatomy and Physiology of the Mammalian Auditory System

Manuel S. Malmierca
Department of Cellular Biology and Pathology, Faculty of Medicine, University of Salamanca, Salamanca, Spain
Auditory Neuroscience Laboratory, Institute for Neuroscience of Castilla y Léon, Salamanca, Spain

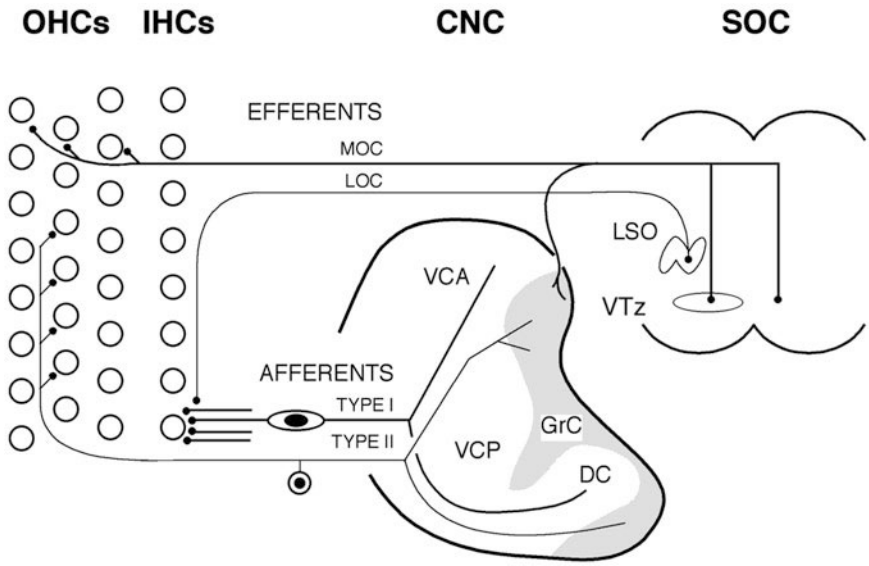
List of Abbreviations

A1	Primary auditory cortex
AC	Auditory cortex
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
AN	Auditory nerve
AVCN	Anteroventral cochlear nucleus
CNC	Cochlear nucleus complex
CNIC	Central nucleus of the inferior colliculus
DAS	Dorsal acoustic stria
DCIC	Dorsal nucleus of the inferior colliculus
DCN	Dorsal cochlear nucleus
DLL	Lateral lemniscus dorsal nucleus
GABA	γ-Aminobutyric acid
IAS	Intermediate acoustic stria
IC	Inferior colliculus
IHC	Inner hair cell
LCIC	Lateral cortex of the inferior colliculus
LOC	Lateral olivocochlear neurons/system
LSO	Lateral superior olive
MGB	Medial geniculate body
MGD	Dorsal division of the medial geniculate body
MGM	Medial division of the medial geniculate body
MGV	Ventral division of the medial geniculate body
MNTB	Medial nucleus of the trapezoid body
MOC	Medial olivocochlear neurons/system
MSO	Medial superior olive
NLL	Nuclei of the lateral lemniscus
NMDA	*N*-Methyl-d-aspartate
OHC	Outer hair cell
PO	Periolivary nuclei
SOC	Superior olivary complex
SPO	Superior paraolivary nucleus
VAS	Ventral acoustic stria
VCN	Ventral cochlear nucleus
VLL	Ventral nucleus of the lateral lemniscus
VNTB	Ventral nucleus of the trapezoid body

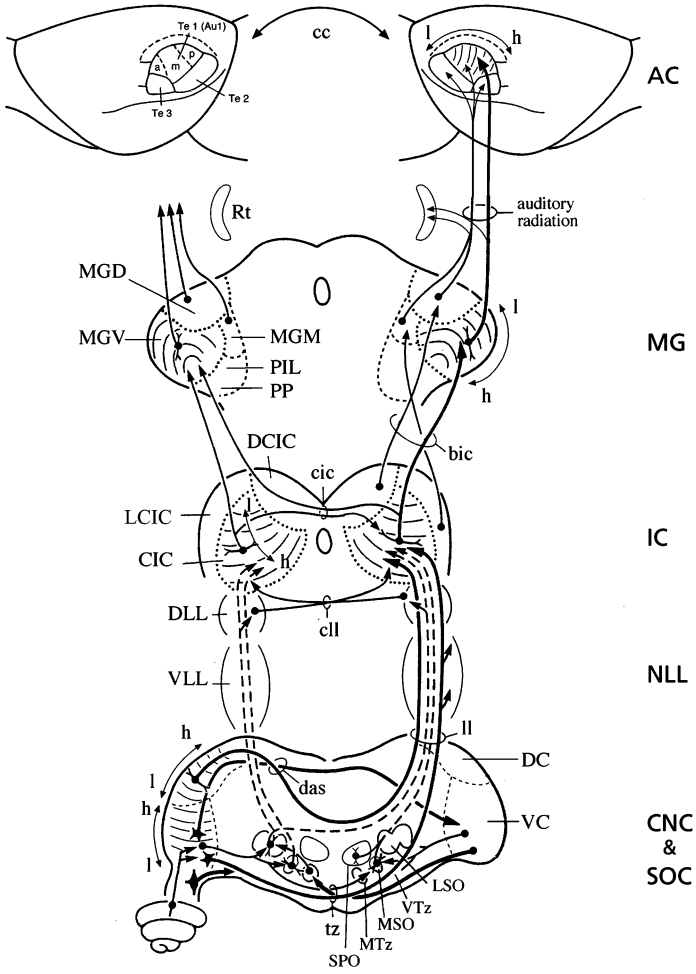
An auditory system is found in all classes of vertebrates, including fish, amphibians, reptiles and birds, and mammals. Although there are important similarities across classes, the system has evolved differently in the different groups. Even within the class of mammals, there are notable specializations, especially in echolocating mammals such as cetaceans and bats. Because one major objective in hearing research is to understand the structure and physiology of the human auditory system, this entry is restricted to an overview of the general plan of organization of the mammalian system. Insights gained from research in animals should aid in identifying the causes of hearing impairments in humans and represent an important step toward developing effective treatments.

The specific auditory stimulus consists of pressure waves arriving at the ear within a certain frequency range. This audible frequency range varies among species (e.g., humans, 0.02–20 kHz; rat, 0.25–70 kHz; mouse, 2–70 kHz; guinea pig, 0.2–45 kHz; and cat, 0.125–60 kHz) (reviewed in Malmierca 2003; Fay 1988). Within their audible range, some species, such as echolocating bats, are tuned to particular frequencies of special importance for their behavior (reviewed in Fay and Popper 1994) and are considered to be "auditory specialists" (Echteler et al. 1994).

Sound waves are transmitted mechanically through the outer and middle ear to the sensory hair cells of the organ of Corti, in the cochlear partition of the inner ear. Auditory nerve fibers transmit information about receptor potentials generated by the sensory hair cells to the brainstem (Fig. 1). In contrast to the minimum of two relay stations between the periphery and cerebral cortex in the visual and somatosensory systems, there is a minimum of three relays in the auditory system (Fig. 2), with several stages of convergence and divergence, and at least seven levels of crossing fibers (Figs. 2 and 3), making the auditory system uniquely complicated. In the first relay center, the cochlear nuclear complex (CNC), signals carried by the cochlear nerve are channeled into a number of parallel ascending tracts (Figs. 1 and 2), each with a particular course and destination and each presumably serving a different function (Fig. 3). Some of these terminate in a collection of nuclei in the pons known as the superior olivary complex (SOC). Ascending auditory tracts from both the CNC and the SOC converge on the inferior colliculus in the midbrain (Figs. 2, 3, and 8). From the midbrain upward, the auditory pathway is often divided into core or "lemniscal" projections with a clear tonotopic organization and belt or "nonlemniscal" projections where tonotopy is less sharp or absent (for review, see Malmierca 2003; Malmierca and Hackett 2010).



Anatomy and Physiology of the Mammalian Auditory System, Fig. 1 Afferent and efferent innervation of the cochlear epithelium. Several type I afferent fibers converge onto single IHCs, while a single type II afferent fibers innervate several OHCs. Type I fibers terminate in the AVCN, and type II fibers terminate on the granule cell regions (*GrC*) and marginal shell areas of the VCN and DCN. The efferent MOC innervate the OHCs, and the efferent LOC innervate IHCs (cf. Figure 19) (Modified after Brown et al. 1988). Abbreviations in the figure: *DC* Dorsal cochlear nucleus, *LTz* Lateral nucleus of the trapezoid body, *VCA* Anteroventral cochlear nucleus, *VCP* Posteroventral cochlear nucleus



Anatomy and Physiology of the Mammalian Auditory System, Fig. 2 Schematic wiring diagram of the ascending auditory pathway (Modified after Brodal 1981, AC is from Herbert et al. 1991). Abbreviations in the figure: *bic* Brachium of the inferior colliculus, *cc* Corpus callosum, *CIC* Central nucleus of the inferior colliculus, *cic* Commissure of the inferior colliculus, *cll* Commissure of the lateral lemniscus (Prosbt), *das* Dorsal acoustic stria, *h* High-frequency region, *DC* Dorsal cochlear nucleus, *l* Low-frequency region, *ll* lateral lemniscus, *LTz* Lateral nucleus of the trapezoid body, *MG* Medial geniculate body, *MTz* Medial nucleus of the trapezoid body, *PIL* Posterior intralaminar nucleus, *PP* Peripeduncular nucleus, *Rt* Auditory sector of the reticular thalamic nucleus, *SPO* Superior paraolivary nucleus, *Te1* Temporal area 1, *Te2* Temporal area 2, *Te3* Temporal area 3, *tz* Trapezoid body (or ventral acoustic stria), *VC* Ventral cochlear nucleus

The Auditory Nerve

The auditory nerve (AN) is made of both *afferent* and *efferent* fibers (for review, see Slepecky 1996). The afferent fibers transmit impulses from the organ of Corti to the cochlear nuclear complex, while the efferent fibers convey impulses from the superior olivary complex to the organ of Corti (Fig. 1). There are two subtypes of afferent fibers: myelinated and unmyelinated fibers (Fig. 1). The myelinated (type I) fibers are relatively thick and originate from bipolar spiral ganglion cells that innervate the inner hair cells (IHCs). The unmyelinated (type II) fibers are thinner and arise from small, pseudounipolar spiral ganglion cells that innervate the outer hair cells (OHCs). About 90–95% of auditory nerve fibers are type I (Fig. 1); each type I fiber terminates on a

single IHC. The type II fibers constitute only about 5% of all auditory nerve fibers (Fig. 1). As opposed to type I, type II fibers are highly branched, and a single fiber forms synapses with many (6–100) OHCs.

Three types of type I fiber have been characterized based on morphological features that correlate with their spontaneous activity and threshold sensitivity (Liberman et al. 1990). Fibers with a low threshold and a high spontaneous firing rate have the larger diameter and terminate on the pillar side of the IHC, whereas fibers with a high threshold and a low spontaneous firing rate are thinner and terminate on the modiolar side.

The efferent fibers of the olivocochlear system belong to the descending auditory pathways (Fig. 1). They can be divided into two groups: the lateral efferent system (lateral olivocochlear system) that innervates auditory nerve fibers near their synapses with IHCs and the medial efferent system (medial olivocochlear system) that innervates the OHCs (Warr 1992).

The Cochlear Nuclear Complex

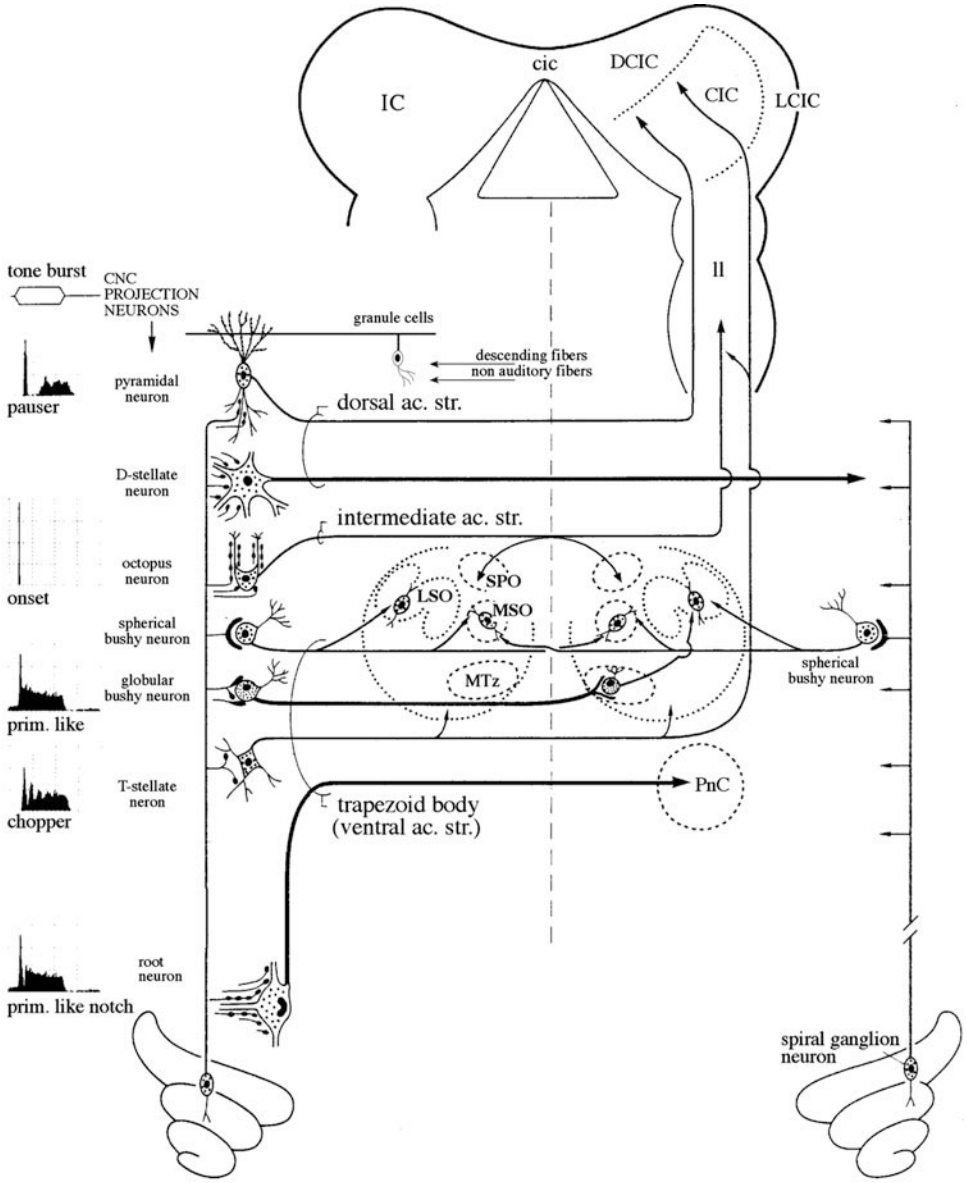
The cochlear nucleus complex is situated laterally and superficially in the brainstem (CNC, Figs. 1–5) and is the first relay center for ascending auditory information. It is the site of termination of all auditory nerve (AN) fibers (Ryugo and Parks 2003). The CNC consists of a ventral cochlear nucleus (VCN) and a dorsal cochlear nucleus (DCN). The VCN is subdivided by the cochlear nerve root into anteroventral (AVCN) and posteroventral (PVCN) parts. The DCN curves around the inferior cerebellar peduncle in the floor of the lateral recess of the fourth ventricle. The axons of CNC projection neurons leave the complex via the three primary pathways to reach higher auditory structures: the dorsal, intermediate, and ventral acoustic striae (DAS, IAS, and VAS, respectively). The VAS is usually referred to as the trapezoid body (Figs. 2 and 3). The projections are largely tonotopically organized, and neurons within an isofrequency lamina of the CNC project to a corresponding isofrequency lamina in higher-order centers. The right and left CNC are also interconnected by fibers of glycinergic commissural neurons as described below (for review, see Cant and Benson 2003). In addition to its ascending inputs from the AN, the CNC receives descending projections from the auditory cortex, the inferior colliculus; the ventral complex of the lateral lemniscus; and the superior olivary complex (see Malmierca 2003). A large proportion of the latter fibers may be inhibitory, glycine and/or GABA being the transmitters, but there are also excitatory descending fibers, e.g., collaterals of the cholinergic olivocochlear bundle (Osen et al. 1984). The CNC also receives some projections from non-auditory brain structures.

Primary Afferents

Each AN fiber bifurcates into an ascending branch, which supplies the AVCN, and a descending branch, which supplies the PVCN and DCN (Fig. 1). The anatomical distribution of the primary fibers forms the basis for the laminar tonotopic organization of the three subnuclei observed electrophysiologically (for review, see Ryugo and Parks 2003). Type I fibers supply all parts of the CNC except the superficial granule cell areas and the molecular layer of the DCN (Fig. 1). Two basic types of terminals are found: large, axosomatic endings called “endbulbs of Held” and small boutons. The endbulbs of Held arise from the ascending branches, while small boutons arise from loosely ramifying collaterals of both ascending and descending branches. The type II fibers innervate areas rich in granule cells and appear to supply the marginal shell of the VCN (Fig. 1) (for review, see Ryugo and Parks 2003).

Ventral Cochlear Nucleus

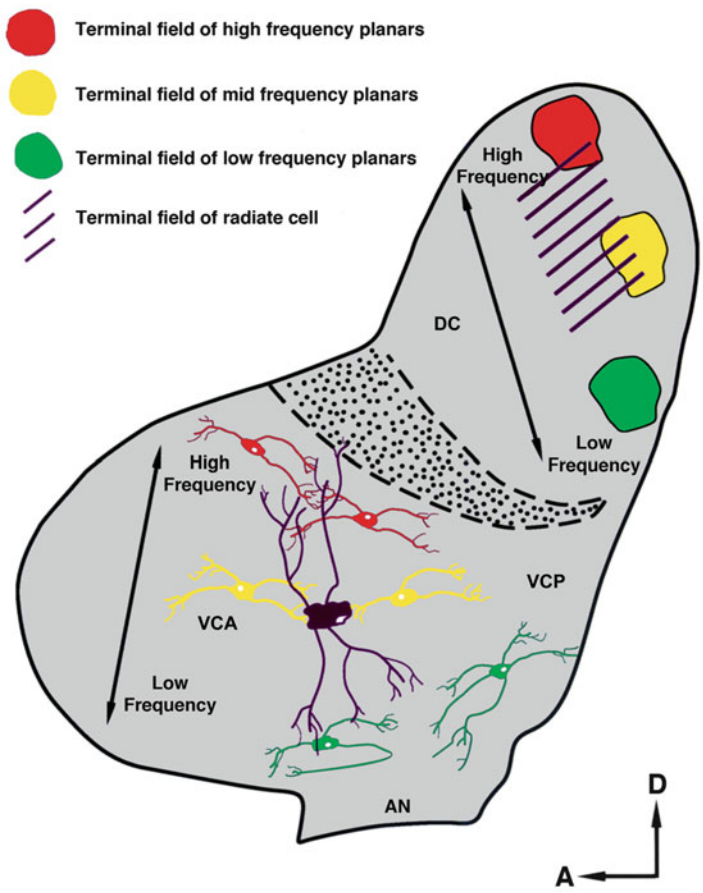
Five main neuronal types are recognized in the VCN based on patterns of Nissl staining, dendritic arborization, and main axonal projections: spherical bushy, globular bushy, octopus, multipolar



Anatomy and Physiology of the Mammalian Auditory System, Fig. 3 Projecting cell types of the CNC and their corresponding physiological responses (Modified after Moore and Osen 1979). For abbreviations, see list. Abbreviations in the figure: *CIC* Central nucleus of the inferior colliculus, *cic* Commissure of the inferior colliculus, *ll* lateral lemniscus, *MTz* Medial nucleus of the trapezoid body, *PnC* Pontine reticular nucleus, caudalis, *SPO* Superior paraolivary nucleus

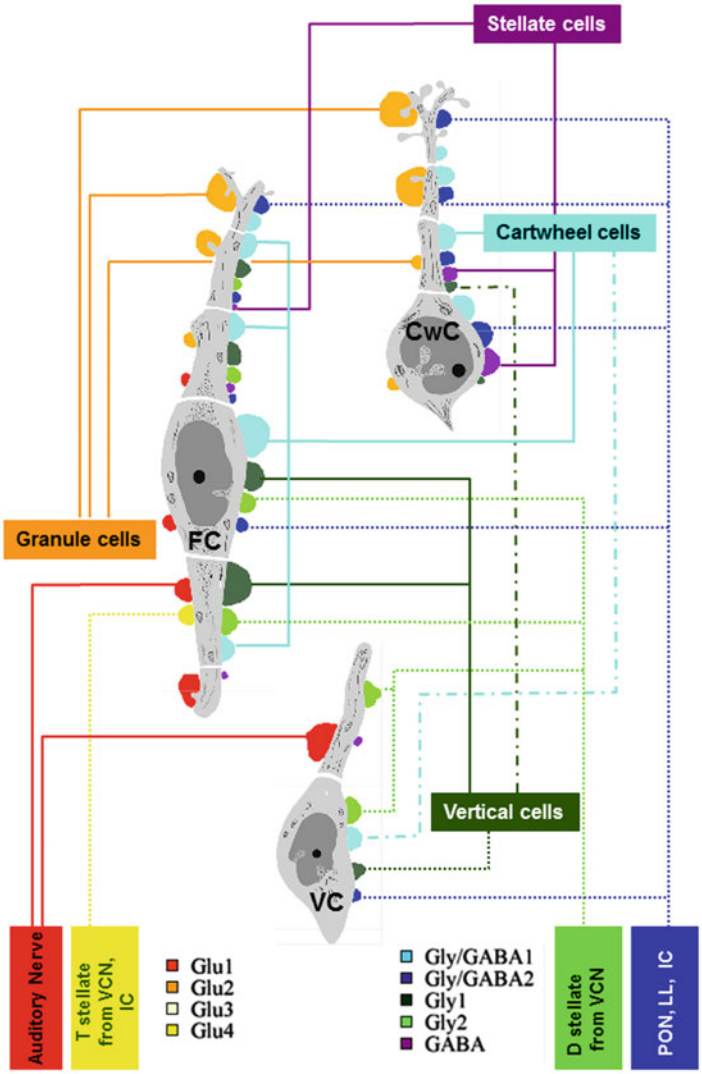
(or stellate), and small cells (Figs. 3–5; Osen 1969; Brawer et al. 1974).

The spherical bushy cells are found rostrally in the AVCN, the globular bushy cells lie centrally on both sides of the nerve root in the caudal AVCN and the rostral PVCN, and the octopus cells are found caudally in the PVCN. *The spherical bushy, globular bushy, and octopus neurons* (Fig. 3) have non-tapering dendrites that end in bushy-like formations, but they differ with regard to the appearance of the terminal bush, the number of root segments, and the relative length of the



Anatomy and Physiology of the Mammalian Auditory System, Fig. 4 Diagram of the planar (T-stellate) and radiate (d-stellate) cells in the VCN and their projections to the DCN. Planar cells have frequency-specific projections, while radiate cells have across frequency projections (Figure kindly provided by Dr. D. K. Ryugo. Reproduced from Doucet and Ryugo 1997). For abbreviations, see list. Abbreviations in the figure: *AN* auditory nerve, *DC* Dorsal cochlear nucleus, *LTz* Lateral nucleus of the trapezoid body, *VCA* Anteroventral cochlear nucleus, *VCP* Posteroventral cochlear nucleus

stem dendrites. The spherical bushy cells receive a small number of large axosomatic terminals, the endbulbs of Held (for review, see Ryugo and Parks 2003), and have the so-called primary-like responses to pure tone stimulation, similar to those of the auditory nerve fibers (Young et al. 1988; Young and Davis 2002). The spherical bushy cells (Fig. 3) project bilaterally to the medial superior olive and to the ipsilateral lateral superior olive (Cant and Benson 2003). Globular bushy cells receive a larger number of distinct inputs than do the spherical bushy cells, including small (or "modified") endbulbs. In response to pure tone stimulation, they exhibit a firing pattern known as "primary-like with notch." The globular bushy cells (Fig. 3) project to the ipsilateral lateral nucleus of the trapezoid body and the contralateral medial nucleus of the trapezoid body. Axons from both globular and spherical bushy cells course ventrally in the trapezoid body. The patterns of their AN input, along with unique cell membrane properties, make these cells capable of transmitting precise temporal information necessary for both high- and low-frequency sound localization (Young et al. 1988; Young and Davis 2002). The octopus cells (Fig. 6) receive small boutons from collaterals arising from the descending branch of the AN. They respond to a tone burst with a single spike and so have been called onset units (Young et al. 1988; Young and Davis 2002). Their main projection is to the superior paraolivary nucleus on both sides and to the contralateral ventral complex of the lateral lemniscus, and their axons course dorsally, looping over the restiform body in the intermediate acoustic stria (Wickesberg and Oertel 1988; for review, see Cant and Benson 2003). Their function is still unclear, but it has been suggested that they encode

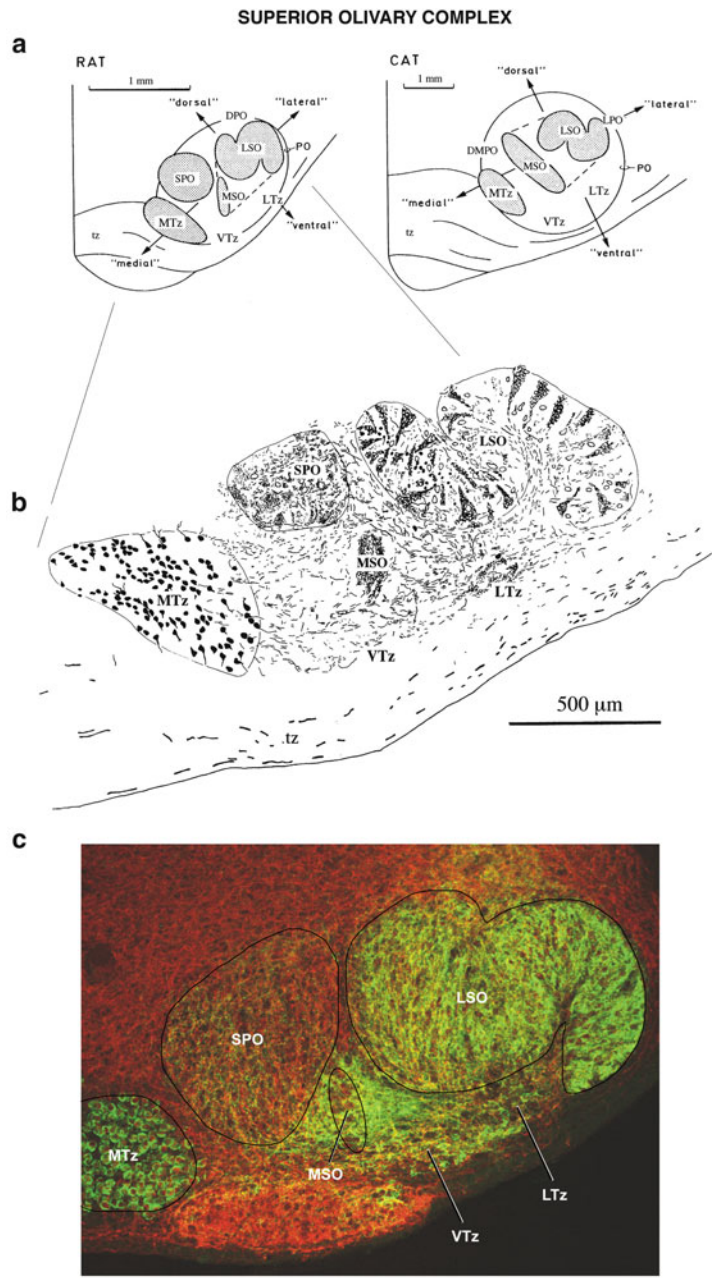


Anatomy and Physiology of the Mammalian Auditory System, Fig. 5 Synaptic endings containing glutamate, glycine, and GABA in the rat dorsal cochlear nucleus. Synaptic endings and of the known (*solid lines*) and putative (*dashed* and *doted lines*) neuronal sources of excitatory and inhibitory endings onto pyramidal (*FC*), vertical (*VC*), and cartwheel (*CWC*) cells (Figure kindly provided by Dr. M. E. Rubio. Reproduced from Rubio and Juiz 2004)

the pitch period in their temporal firing patterns (Oertel 1999).

The *multipolar* and *small* cells are present throughout the VCN (Figs. 3 and 4). The small cells are most abundant around the peripheral margins of the nucleus deep to the superficial granule cell layer. A large collection of small cells located dorsolaterally in a superficial location forms the *small cell cap* of the VCN (Osen 1969). This is particularly conspicuous in cat. *The multipolar cells* (Figs. 3 and 6) possess moderately branched, tapering dendrites which are contacted by small boutons from many primary afferent fibers. Two types of multipolar cell have been described (Fig. 3): type I, also referred to as T-stellate (mouse) and planar (rat), and type II, also referred to as d-stellate (mouse) or radiate (rat).

Multipolar type I cells (Figs. 3 and 4) have oriented dendritic arbors and project to the periolivary region of the superior olivary complex via the trapezoid body, the nuclei of the lateral lemniscus, and the central nucleus of the inferior colliculus through the lateral lemniscus (Adams 1979; Cant and Benson 2003; Malmierca et al. 2005). They also give rise to frequency-specific collaterals within the VCN and DCN (Lorente de Nó 1981). These multipolar neurons exhibit a



Anatomy and Physiology of the Mammalian Auditory System, Fig. 6 (**a**) Comparison of the superior olivary complex in the rat and cat (Redrawn after Osen et al. 1984). Note the relative size of of the LSO–MSO in the two species and the existence of a distinct SPO in the rat. (**b**), Camera lucida drawing of a section showing calbindin-positive neurons in the MTz and processes in the rat SOC (Redrawn after Friauf 1993). (**c**) Confocal image illustrating VGLUT1-ir in the SOC of adult rats. VGLUT1-ir is *green*, and MAP 2-ir is *red*. For abbreviations, see list (Data from Blaesse et al. 2005. Figure kindly provided by Dr. E. Friauf). Abbreviations in the figure: *DMPO* Dorsomedial periolivary region, *LPO* Lateral periolivary zone, *LTz* Lateral nucleus of the trapezoid body, *MTz* Medial nucleus of the trapezoid body, *PO* Periolivary regions, *SPO* Superior paraolivary nucleus, *tz* Trapezoid body (or ventral acoustic stria); *VTz* Ventral nucleus of the trapezoid body

"chopper" response to tone bursts with regularly repeated firing. They may be specialized for conveying frequency-specific excitatory information about complex acoustic stimuli such as speech.

Multipolar type II cells (Figs. 3 and 4), also known as commissural neurons, have non-oriented dendritic arbors. They project to the contralateral CNC via the dorsal acoustic stria (Smith and Rhode 1989; Oertel et al. 1990; Doucet and Ryugo 1997). They also emit extensive (probably broadly tuned) collaterals within the ipsilateral VCN and DCN. They are the only known inhibitory (glycinergic) projection neurons of the CNC (Osen et al. 1990; Doucet et al. 1999)

and respond to pure tone stimulation with an "on-chop" pattern (Smith and Rhode 1989).

The small cells are abundant in the marginal shell of the VCN, which is composed of the "granule cell layer" and the subjacent "cap area". The granule cell layer is continuous over the free surface of the CNC and forms a lamina partly separating the VCN and DCN (Mugnaini et al. 1980a, b; for review, see Cant and Benson 2003). In the DCN, the granule cell layer is covered superficially by a molecular layer. The granule cell axons project as parallel fibers (Fig. 5) to the molecular DCN layer (Mugnaini et al. 1980a, b). The cap area is small but still distinguishable due to its contingent of small cells, many of which show glycine- and/or GABA-like immunoreactivity (for review, see Cant and Benson 2003). The cap is supplied by both type I and type II fibers (Fig. 1), and at least in cat, nearly all type I auditory nerve fibers that innervate the cap have low spontaneous rates (for review, see Ryugo and Parks 2003). The marginal shell also receives descending input. Its cells show a wide dynamic range (Ye et al. 2000) and have a large diversity of projections (Adams 1979; Malmierca et al. 2002, 2005; Ye et al. 2000). The available electrophysiological studies suggest that they form part of a feedback gain control system made up of the cochlea, cochlear nuclear complex, medial olivocochlear system, and outer hair cells (Ye et al. 2000).

Cochlear Root Neurons (Fig. 3). The CNC of rodents contains a population of large cells scattered in the cochlear nerve root, between the main body of the VCN and the glial Schwann-cell border of the AN. It has been suggested that these root neurons participate in the acoustic startle reflex (Sinex et al. 2001).

Dorsal Cochlear Nucleus

The DCN varies from being markedly laminated in rodents and carnivores, where it resembles the cerebellar cortex, to appearing nonlaminated in humans (but see Rubio et al. 2008) and some bat species; it is virtually absent in some cetaceans. The three superficial layers of the DCN are related to the morphology of the principal pyramidal (fusiform) cells (Figs. 3 and 5). The spiny apical dendritic arbor of pyramidal cells occupies layer 1 together with granule cell axons and several other types of interneurons (Fig. 5, see below). Pyramidal cell bodies define layer 2, and their aspinous basal dendritic arbors comprise layer 3. Pyramidal cell dendritic arbors are flattened across the long, frequency gradient axis of the DCN (see, e.g., Cant and Benson 2003; Malmierca 2003). The highest degree of flatness and mutually parallel orientation is found in the basal arbor, which is supplied by primary afferent fibers in a strictly tonotopic manner.

The pyramidal cells are the main projection neurons of the DCN, supplying fibers to the contralateral IC via the DAS (Fig. 3). In addition, some have a direct projection to the medial division of the medial geniculate body (Malmierca et al. 2002). The deepest layer of the DCN contains two categories of cells based on their size, the giant cells which project to the contralateral IC through the DAS and smaller glycinergic tuberculoventral interneurons (Fig. 5). Pyramidal and giant cell excitatory responses are more strongly influenced by their inhibitory inputs than are those of other projection neurons in the CNC and have been classified as types III and IV (Oertel and Young 2004). The type IV units have been found to be sensitive to spectral notches created by the pinna, which may be important cues for localizing sounds. DCN projection neurons receive and respond not only to auditory information but also to somatosensory inputs from muscle proprioceptors in and around the pinna. This innervation has led to speculation that the DCN may be involved in coordinating pinna orientation with localization cues found in the different spectra of sounds located at different points in space. In fact, bilateral lesions of the DAS in cats result in reduced accuracy in head orientation responses to broadband sounds, particularly in elevation (reviewed in Young and Davis 2002).

DCN possesses a large number of interneurons that may be divided into two systems: the tuberculoventral system and the granule cell system (Fig. 5). The *tuberculoventral system*

reciprocally interconnects the DCN and VCN. It contains both frequency-specific and diffuse projections (Malmierca 2003). The frequency-specific projection from the DCN to the VCN originates from small interneurons, a subset of the glycinergic "vertical cells" (Fig. 5). A separate set of vertical cells with only local collaterals contain both GABA and glycine, the relative amounts of which vary among species. They are located between the basal pyramidal cell dendrites in layer 3, have flattened dendritic arbors, and provide the DCN and VCN with a tonotopically organized inhibition. One projection from the VCN to the DCN is made up of the collaterals of type I multipolar cells described above and probably is excitatory and tonotopic. An inhibitory projection arises from axonal collaterals of the glycinergic commissural (type II) cells. The vertical cells of the DCN provide inhibition over a narrow frequency range, whereas the on-chop, type II stellate cells generate inhibition over a wide frequency range.

The granule cell system includes two types of excitatory cells: granule cells and unipolar brush cells and three types of inhibitory cells: the GABAergic Golgi and stellate cells and the glycinergic cartwheel cells (Fig. 5; Oertel and Young 2004; Rubio and Juiz 2004). The granule cells receive direct excitatory input from many sources including the somatosensory system and inhibitory inputs via the Golgi cells. The granule cell axons contribute parallel fibers to the molecular layer. The unipolar brush cells seem to represent a device for feedforward excitation to the mossy fiber pathways, while the stellate cells and cartwheel cells provide feedforward inhibition to the pyramidal cells (Rubio and Juiz 2004; Fig. 5).

The Superior Olivary Complex

The superior olivary complex (SOC) comprises a group of nuclei in the caudal pons (Figs. 2 and 6). Three main nuclei are consistently identified: the lateral superior olive (LSO), the medial superior olive (MSO), and the medial nucleus of the trapezoid body (MNTB). These three nuclei are surrounded by more diffusely organized cellular areas, collectively referred to as the periolivary region (PO) (Adams 1983; Osen et al. 1984; Schofield and Cant 1991). The LSO and MNTB are well developed in both the rat and cat, while they are relatively small in human. In contrast, the MSO is small in the rat but is prominent in both the cat and human (Fig. 6). These differences appear to be related to the ability to use specific frequency cues for directional hearing. The MSO extracts the information about interaural timing differences that is available in low-frequency sounds. Together, the LSO and MNTB detect interaural intensity differences generated by high-frequency sounds.

Lateral Superior Olive

The LSO consists of layers of flattened multipolar neurons (principal cells) with their dendrites oriented perpendicular to its long axis, which is curved into an S-shape. The LSO is tonotopically organized with low frequencies represented laterally and high frequencies, medially. In addition to the principal cells, other, less abundant, neuronal types are also present (Rietzel and Friauf 1998). In some species, neurons of the lateral olivocochlear system lie within the LSO.

The LSO receives direct input from the AVCN on the ipsilateral side and indirect input from the AVCN on the contralateral side (Fig. 6). The ipsilateral input derives from spherical bushy cells and is excitatory. Globular bushy cells in the contralateral AVCN project to the MNTB ipsilateral to the LSO. The MNTB, in turn, sends an inhibitory (glycinergic) input to the LSO (Fig. 6). Multipolar type I neurons from the AVCN also innervate the LSO (reviewed in Helfert and Aschoff 1997; Thompson and Schofield 2000).

The LSO projects to the central nucleus of the inferior colliculus bilaterally (Figs. 2 and 6). Most of the ipsilaterally projecting cells are inhibitory (glycinergic), while the contralaterally projecting cells are glycine-negative and presumably excitatory. The LSO also innervates the dorsal nucleus of the lateral lemniscus bilaterally (Fig. 2).

Medial Nucleus of the Trapezoid Body

Cells of the MNTB are situated among fascicles of fibers in the trapezoid body (Figs. 2 and 6). The principal cells resemble the globular bushy cells of the AVCN, whereas non-principal (marginal) cells have a multipolar appearance (Fig. 6; Morest 1968; Banks and Smith 1992; Sommer et al. 1993).

The MNTB receives input from the globular bushy cells in the contralateral VCN (Fig. 2). These cells give rise to thick axons that terminate on the principal cells in the MNTB as large axosomatic calyces of Held (1893) in a one-to-one relationship. These calyces provide a fast and secure relay of information from the CNC to the MNTB and from there to the LSO. They constitute the largest synaptic terminals in the mammalian brain. The responses of cells in the MNTB to acoustic stimuli show a sharp onset spike that is also characteristic of the globular bushy cells.

In addition to its projection to the LSO, the MNTB projects to the ipsilateral MSO as well as other parts of the ipsilateral SOC and the dorsal portion of the ventral complex of the lateral lemniscus, providing a source of widespread glycinergic inhibition.

The microcircuitry and neurochemistry of the LSO and MNTB suggest that a major function of the MNTB is to transform excitatory contralateral input into ipsilateral inhibition, allowing the LSO to faithfully encode interaural intensity differences in the high-frequency range of audition (for review, see Kopp-Scheinpflug et al. 2008).

Medial Superior Olive

The MSO lies between the LSO and the MNTB (Fig. 6) and is populated by two types of cells: principal and non-principal or marginal cells. The bipolar principal cells are organized into a transversely oriented row, with their dendrites extending in the medial and lateral directions. The multipolar non-principal cells are scattered among these dendrites and are much less numerous (Smith 1995). The MSO is tonotopically organized with low-frequency tones represented dorsally and high-frequency tones ventrally, although most of the nucleus appears to be devoted to low frequencies (Guinan et al. 1972).

The MSO receives direct, excitatory input from spherical bushy cells in the AVCN bilaterally (Fig. 6; for review, see Malmierca 2003), but these inputs remain segregated on the cell surface, such that the lateral dendrites receive input from the ipsilateral side, while the medial dendrites receive input from the contralateral side (Smith 1995). A rostrocaudal organization of the axons is reminiscent of the required input configuration in the Jeffress model for sound localization. The direct bilateral input suggests that the MSO neurons are ideally suited to measure interaural phase or time differences (Joris et al. 1998).

The MSO also receives inhibitory inputs, mostly glycinergic, from the medial and lateral nuclei of the trapezoid body on the same side. The latter may also provide GABAergic inhibitory input (Smith 1995). The principal cells project to the inferior colliculus and to the ipsilateral dorsal nucleus of the lateral lemniscus (Fig. 8).

Superior Paraolivary Nucleus

A fourth distinct nucleus in the SOC of rodents, the so-called superior paraolivary nucleus (SPO), is found in the dorsomedial part of the complex (Fig. 6; Osen et al. 1984; Schofield and Cant 1991; Schofield 1995). It consists of GABAergic multipolar cells, which are the largest in the SOC, and receives inputs from octopus and multipolar cells in the contralateral VCN, from multipolar cells in the ipsilateral VCN, and a substantial glycinergic input from the MNTB on the same side. SPO projects to the ipsilateral IC (Schofield 1995) and may represent a hyperdevelopment of periolivary cells with a similar projection, present in smaller numbers in other mammals (Adams 1983). It appears that the SPO neurons are well suited for the analysis of temporal features of

complex sounds and stimulus features across broad frequency ranges (Dehmel et al. 2002; Kulesza et al. 2003).

Periolivary Nuclei

The PO contains several distinct types of neurons with different projection patterns (Adams 1983; Osen et al. 1984). Neurons in the PO areas receive input from VCN bilaterally, the lateral areas, from the ipsilateral side, and the medial areas, from both sides (Fig. 6). Certain parts of the PO also receive input from the ipsilateral MNTB (probably inhibitory), from the ipsilateral inferior colliculus (probably excitatory), and from the dorsal nucleus of the lateral lemniscus (probably inhibitory).

PO cells project either to the cochlea, the CNC, or the IC, but individual cells do not appear to project to all three structures (Adams 1983). The ventral nucleus of the trapezoid body (VNTB) is a PO region situated ventral to the MNTB and is of particular interest because it may be involved in the activation of the olivocochlear neurons (Rajan 1990). It is strategically situated at the intersection of ascending projections from the CNC and descending projections from the IC. The VNTB receives major afferent projections from the contralateral VCN, the ipsilateral PVCN (Smith et al. 1991; Thompson and Schofield 2000), and the marginal shell in the AVCN from both sides (Ye et al. 2000). It is the major target in the SOC for the descending projections from the IC.

The Nuclei of the Lateral Lemniscus

The nuclei of the lateral lemniscus (NLL) is made is of two distinct functional systems (Fig. 2), a monaural *ventral* (*VLL*) and a binaural *dorsal (DLL)* system (Covey and Casseday 1991). There are connectional, neurochemical, and physiological properties which are unique to each system (Covey and Casseday 1991; Malmierca et al. 1998).

The Ventral Nucleus of the Lateral Lemniscus: The Monaural System

The VLL consists of groups of neurons embedded within the part of the lateral lemniscus, located between the SOC and DLL (Fig. 2). It receives its inputs mainly from the contralateral ear via the contralateral VCN and ipsilateral MNTB (Fig. 2). VLL neurons exhibit a variety of shapes and sizes in Nissl-stained sections, and most of them project to the ipsilateral IC. The majority of cells in the ventral part of the complex are glycine and/or GABA (Riquelme et al. 2001).

In vitro, some VLL neurons show an onset firing pattern and a nonlinear current-voltage relationship, while others exhibit a linear current-voltage relationship and other firing patterns (Wu 1999; Zhao and Wu (2001)). Similar differences have also been found in in vivo studies (Zhang and Kelly 2006a, b). The VLL neurons are suitable for encoding temporal events (Covey and Casseday 1991).

The Dorsal Nucleus of the Lateral Lemniscus: The Binaural System

The DLL receives input from both ears, and it projects both to ICs and to its counterpart on the opposite side through the commissure of Probst (Fig. 2). The DLL plays an important role in functions dependent on binaural processing such as sound localization.

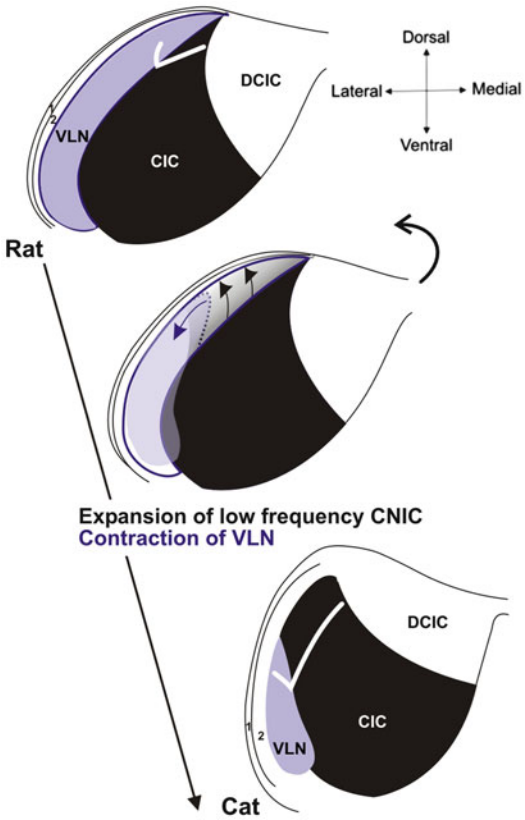
Several neuronal types have been described, depending on the species and the criteria used for cell classification. Regardless of the morphological type, all cells have similar membrane properties with a sustained series of regular action potentials produced by the injection of positive current.

Generally speaking, the DLL receives collaterals from afferent fibers that also innervate the IC (Fig. 8). Thus, DLL receives contralateral inputs from the ventral cochlear nucleus and DLL, ipsilateral input from the medial superior olive, superior paraolivary nucleus and VLL, and bilateral inputs from the lateral superior olive. The DLL projection to the IC is laminar and bilateral, with a

predominant projection to the contralateral IC. Most DLL cells are GABAergic and therefore have an inhibitory influence on the IC.

The Inferior Colliculus

The inferior colliculus (IC) is the principal auditory nucleus in the midbrain and is characterized by a massive convergence of inputs from lower and higher auditory centers as well as from nonauditory structures (Figs. 2 and 7; Irvine 1992; Malmierca 2003; Casseday et al. 2002; Loftus et al. 2008). The IC is divided into a central nucleus (CNIC) surrounded by cortical regions (Fig. 7). These collicular cortices include a dorsal cortex (DCIC) that covers the CNIC dorsally and caudally, a lateral cortex (LCIC) that covers it laterally, and a rostral cortex (RCIC) that covers it rostrally (Loftus et al. 2008; Fig. 8).



Anatomy and Physiology of the Mammalian Auditory System, Fig. 7 Subdivisions of the inferior colliculus in the rat and cat. The low-frequency representation in the central nucleus (*CIC*) is expanded in cat and contracted in rats, while the size of the ventrolateral nucleus (*VLN*) (layer 3 of LCIC) is expanded in rat and contracted in cats. (Redrawn from Loftus et al. 2008)

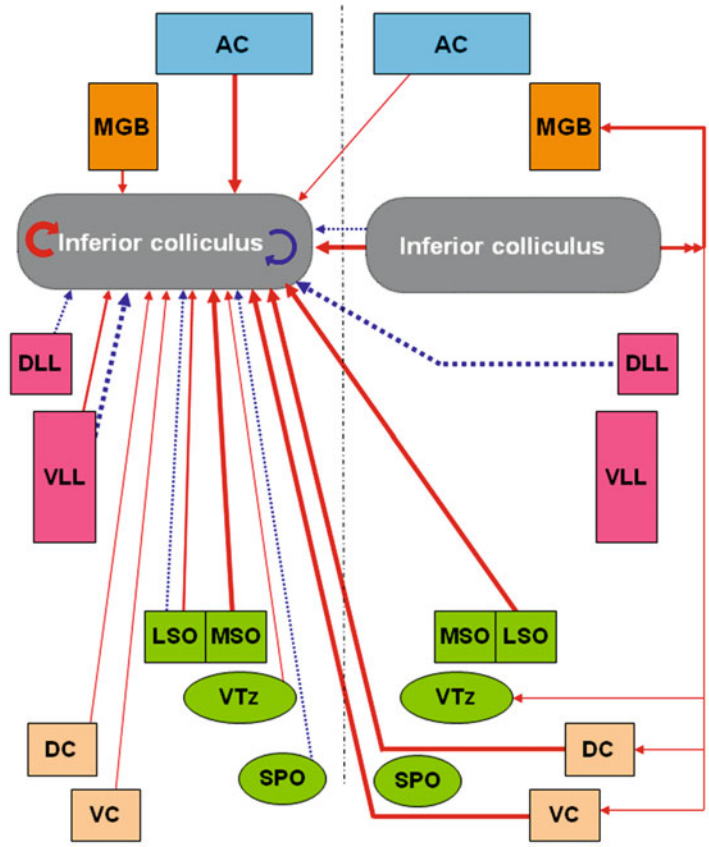
Neurons in the CNIC tend to be most strongly influenced by lower auditory centers, while neurons in the DCIC and RCIC tend to be most strongly influenced by the descending pathways and commissural inputs. The LCIC (and probably the RCIC) receives both auditory and nonauditory (e.g., somatosensory) inputs.

In addition, to its intrinsic and commissural fiber systems (Malmierca et al. 1995; Saldaña and Merchán 1992), the IC provides the major ascending projections to the MGB as well as descending projections to the SOC and the CNC (Oliver et al. 1999; Malmierca et al. 1996; Peruzzi et al. 1997; Ito and Oliver 2012; Cant and Benson 2007).

The Central Nucleus of the Inferior Colliculus

The CNIC is distinguished by layers of cells and fibers organized into "fibrodendritic laminae" (Oliver and Morest 1984; Faye-lund and Osen 1985; Malmierca et al. 1993). These consist of a parallel organization of afferent lemniscal fibers and neurons with flattened dendritic arbors and constitute the structural basis for the tonotopic organization of the IC (Schreiner and Langner 1997; Malmierca et al. 2008; Fig. 9).

The CNIC is composed of two main neuronal types: disk-shaped or flat neurons and stellate or less flat neurons (Oliver and Morest 1984; Malmierca et al. 1993). These neurons differ in several respects, including the thickness, branching pattern and orientation of their dendritic arbor, and their location with regard to the laminae (Fig. 10). The flat neurons lie parallel to the ascending fibers and form laminae that are mostly one cell thick, about 40–70 μm (Faye-Lund and Osen 1985; Malmierca et al. 1993). These laminae are separated by interlaminar compartments that are populated by the stellate neurons. IC neurons show complex functional

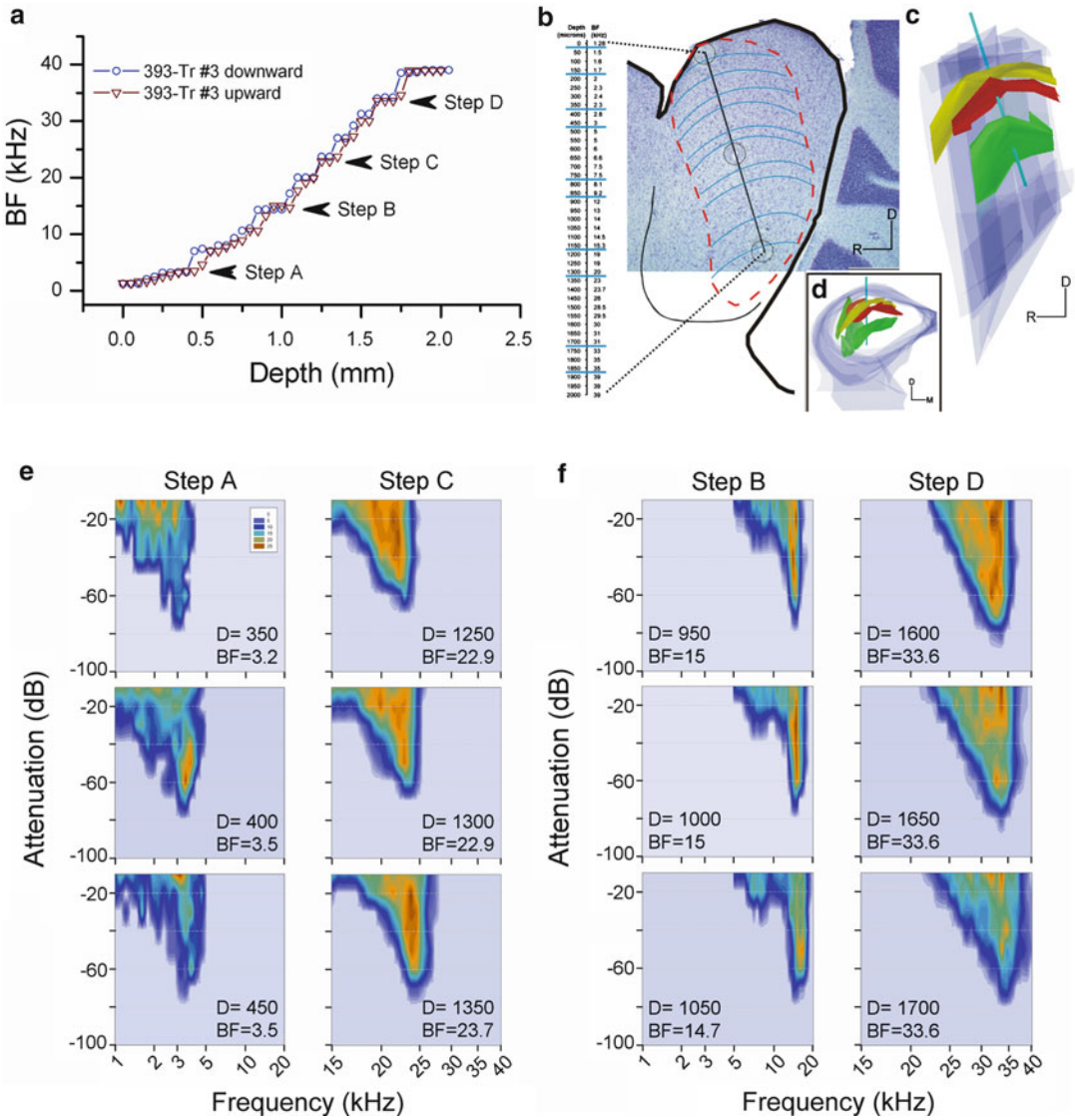


Anatomy and Physiology of the Mammalian Auditory System, Fig. 8 Summary of major sources of inputs to the inferior colliculus on one side of the brain (*left*) from auditory brainstem (*ascending* inputs) and auditory cortex (*descending* inputs). Connection strength is denoted by the thickness of the *arrows*. *Red lines* represent excitatory connections, and *blue/dashed lines* represent inhibitory connections. Abbreviations in the figure: *DC* Dorsal cochlear nucleus, *LSO* Lateral superior olive, *MSO* Medial superior olive, *SPO* Superior paraolivary nucleus, *VC* Ventral cochlear nucleus, *VTz* Ventral nucleus of the trapezoid body

properties, and there is no simple correlation between physiological response properties and morphological classifications (Sivaramakrishnan and Oliver 2001).

The CNIC receives *ascending afferent projections* (Fig. 8) that originate in the AVCN, PVCN, and DCN bilaterally (although mainly contralaterally); the VLL and MSO ipsilaterally; and the DLL and LSO bilaterally (Malmierca 2003; Cant and Benson 2006). These inputs are tonotopically organized and tend to show a "banded" pattern of projections with dense bands (about 200 μm thick) parallel to the isofrequency fibrodendritic laminae. The terminal fields of the various ascending projections may also vary in their distribution along the main axis of the IC. For example, in the dorsomedial parts of the laminae, the axons from the DCN do not overlap with afferents from the LSO. The CNIC projects to the ventral division of the MGB in a tonotopic manner (Fig. 10), largely to the ipsilateral side but also with a crossed component. A majority of neurons use glutamate as the neurotransmitter, but about a quarter of the cells in the CNIC are GABAergic (Merchán et al. 2005), and some of these project to the MGB, making short-latency, monosynaptic inputs (Peruzzi et al. 1997).

IC neurons possess NMDA and AMPA receptors as well as $GABA_A$, $GABA_B$, and glycine receptors (reviewed in Kelly and Caspary 2005). Studies using microiontophoresis in vivo have demonstrated that the AMPA receptors regulate the onset response. Both AMPA and NMDA are involved in the maintenance of the response for the duration of the stimulus, and both GABA and glycine inhibit IC neurons and shape different temporal, spectral, and binaural properties (e.g., LeBeau et al. 2001; reviewed in Kelly and Caspary 2005). Neurons in the IC show a variety



Anatomy and Physiology of the Mammalian Auditory System, Fig. 9 Summary of the tonotopic organization of the inferior colliculus. (**a**) A single electrode penetration *downward* (*blue trace*) and *upward* (*red trace*) along the same electrode track (*Tr*) along the CIC in which FRA obtained from multiunit clusters were recorded at every 50 μm. (**b**) Frequency representation in the CIC obtained in a sagittal section after Nissl staining showing an electrode track (*black line*), three electrolytic lesions (*circles*), and best frequency recorded at 50 μm intervals. Frequency increases as a function of depth in a stepwise manner. (**c**) 3-D reconstruction of three axonal laminae (*yellow*, 1.7 kHz lamina; *red*, 1.8 kHz; *green*, 4.5 kHz) (Data from Malmierca et al. 2008). (**d**) A frontal view of the same 3-D reconstruction seen in (**c**) after 90° rotation. (**e** and **f**) FRAs recorded at every 50 μm in the electrode penetration illustrated in (**a**) at four different steps (*step A–step D*). *D* Depth. (Redrawn from Malmierca et al. 2008)

of frequency response areas that include both V-shaped and non-V shaped types (LeBeau et al. 2001; Palombi and Caspary 1996). Binaural responses are generally classified as suppression, summation, or mixed. The laminae of the CNIC contain a highly organized representation of both spectral and temporal parameters (for review, see Schreiner and Langner 1988).

The Dorsal Cortex of the Inferior Colliculus

The DCIC covers the dorsomedial and caudal aspects of the CNIC (Figs. 7 and 9b) and is made of three layers (Faye-Lund and Osen 1985). Layer 1 (the most superficial layer) is a thin fibrocellular capsule that wraps the entire surface of the IC including the LCIC. It contains scattered, small, flattened neurons. Layer 2 is slightly thicker and contains mostly small- and medium-sized multipolar neurons. Layer 3 includes a heterogeneous group of multipolar neurons that vary in soma size and shape and in their dendritic orientation. A distinct feature of neurons in DCIC (and also LCIC) is that they contain the neuromodulator nitric oxide, a fact that may explain some of the long-term potentiation and neuronal plasticity observed in some IC neurons (Zhang and Wu 2000).

The ascending input from lower auditory centers to the CNIC encroaches upon the DCIC, as do the intrinsic projections (Saldaña and Merchán 1992; Malmierca et al. 1995). In addition, the DCIC receives descending input bilaterally that originates largely from the primary AC. This projection may generate long-lasting changes in the neuronal responses of the IC. The DCIC projects to the dorsal division of the MGB (Winer 1985).

The Lateral Cortex of the Inferior Colliculus

The LCIC covers the CNIC laterally (Fig. 7) and ventrally (Loftus et al. 2008) and is made up of three layers (Faye-Lund and Osen 1985; Loftus et al. 2008). Layer 1 is a continuation of the fibrodendritic capsule of the DCIC. Layer 2 is composed of small- and medium-sized neurons that are partly aggregated into dense clusters rich in acetylcholinesterase and GABA. Layer 3 of the ECIC constitutes its largest part and appears to continue also into the non-stratified rostral cortex. In addition to auditory inputs, the LCIC receives somatosensory input from spinal cord, dorsal column nuclei, and spinal trigeminal nuclei in the cat (Zhou and Shore 2006). Neurons in the LCIC appear to have relatively broad somatosensory receptive fields in addition to auditory responses, which are also broadly tuned (Aitkin et al. 1978). The multisensory integration in the LCIC mirrors similar types of function at higher levels of the "extralemniscal" auditory pathway. Although the functions of the LCIC are not known, it may be a major source of binaural cues for gaze control in the superior colliculus (King et al. 1998), and it is very likely that LCIC plays an important role in providing auditory input to visual-motor areas that direct the head and eye movements involved in gaze initiation. A second role of the LCIC may be in multisensory integration, distinct from the "classical" auditory role of the central nucleus. Finally, the CNIC and LCIC differentially activate the medial and lateral olivocochlear system (Ota et al. 2004).

The Rostral Cortex of the Inferior Colliculus

The rostral cortex (RCIC) is adjacent to the anterior CNIC (Fig. 9b) and includes the largest multipolar cells in the IC. There are also small- and medium-sized multipolar neurons (Faye-Lund and Osen 1985; Malmierca et al. 1993). Like the LCIC and DCIC, the RCIC receives input from the cerebral cortex as well as from nonauditory structures (for review, see Malmierca 2003). The RCIC projects to the dorsal and medial divisions of the MG (Fig. 2). The functional role of RCIC neurons is still unknown, although recent studies have demonstrated that many neurons in this region (and in other cortical regions) may be specialized for detecting novel sounds (Malmierca et al. 2009).

The Medial Geniculate Body

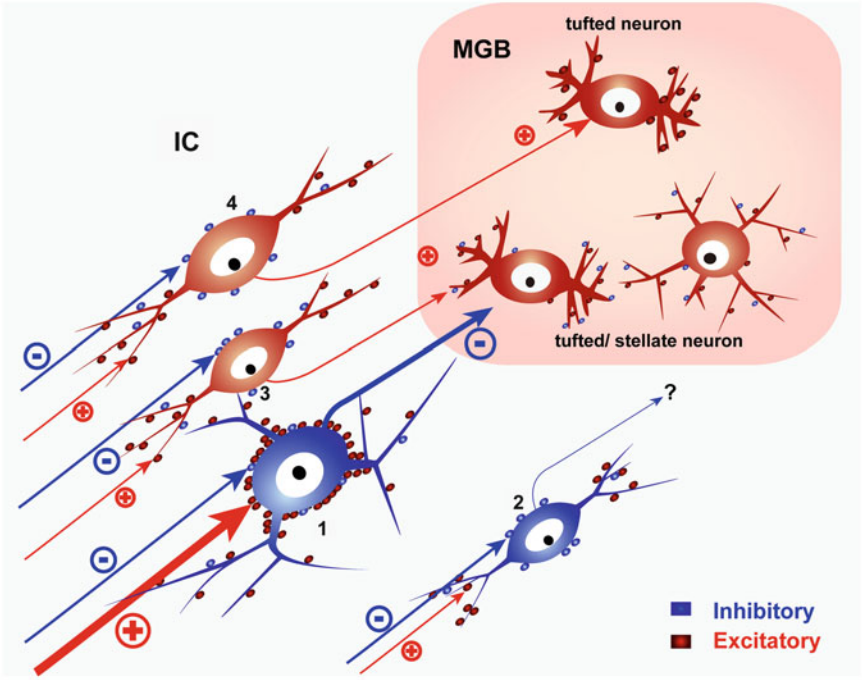
The principal auditory nucleus in the thalamus is the medial geniculate body (MGB), which is also the last opportunity for subcortical processing in the ascending auditory pathways (Fig. 2). The MGB is divided into three major divisions named relative to their locations within the

complex: the ventral (MGV), dorsal (MGD), and medial (or magnocellular) (MGM) divisions (Figs. 11 and 12). Additional subdivisions are also recognized in some species, usually representing smaller domains within each of the major subdivisions. Physiological studies have shown that the caudal part of the reticular thalamic nucleus (Fig. 2) is also a part of the auditory pathway (Cotillon et al. 1999; Yu et al. 2009). It provides the MGB with an inhibitory GABAergic input (Bartlett and Smith 1999; Yu et al. 2009).

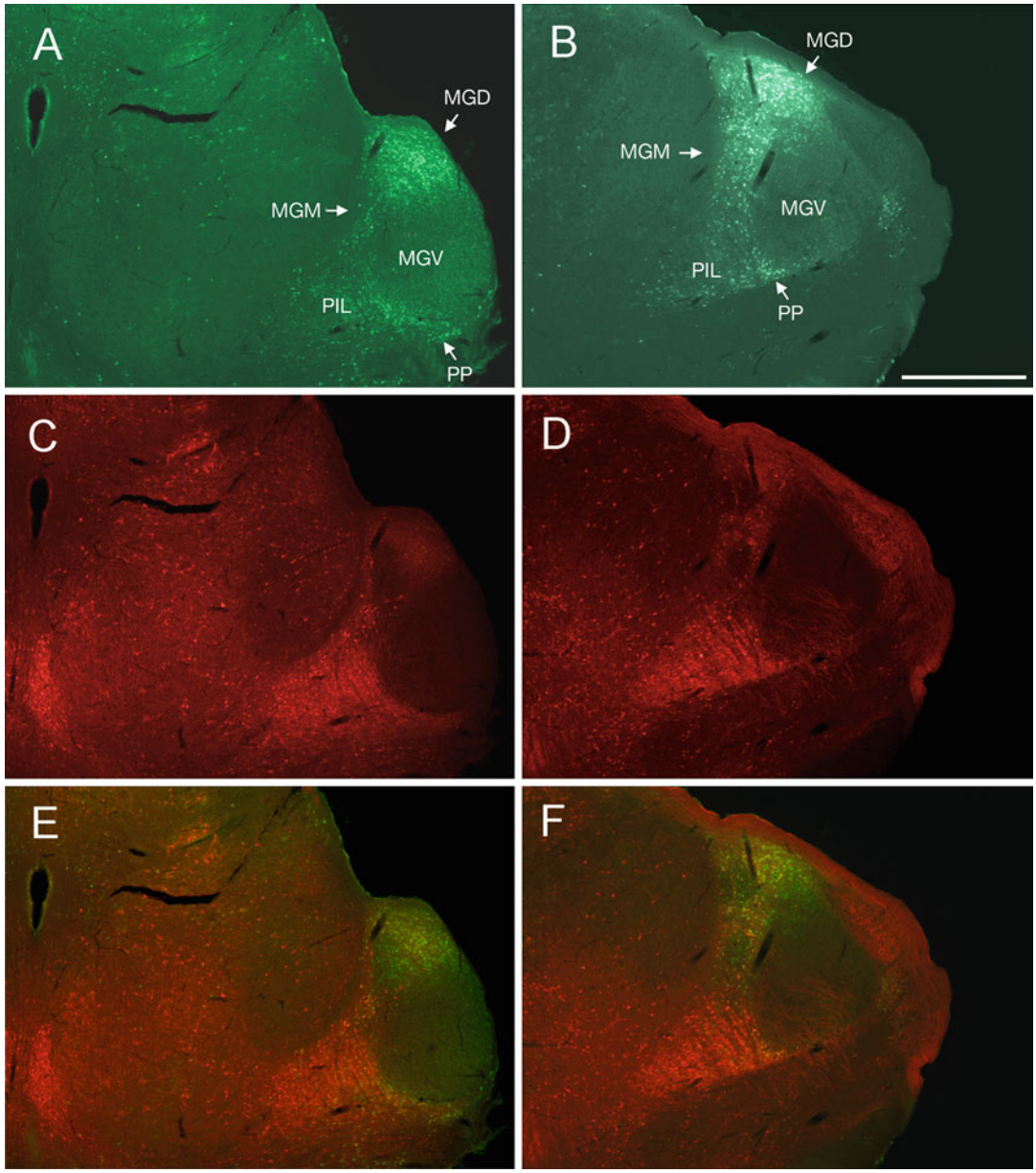
The main source of ascending projections to the MGB is the IC (Figs. 2 and 10), but other inputs include the thalamic reticular nucleus and auditory subcollicular nuclei, including the SOC, NLL, and CNC, challenging the idea that the IC is an obligatory relay station (Malmierca et al. 2002). Ascending fibers enter the structure medially through the brachium of the IC and terminate among neurons in each subdivision. Connections with auditory cortex pass through the posterior limb of the internal capsule, and, for the most part, these connections are reciprocal.

The Ventral Division of the Medial Geniculate Complex

The MGV contains bi-tufted principal neurons that tend to be arranged in rows with oriented dendritic fields (Fig. 12a–c; Winer et al. 1999). These neurons are glutamatergic, and most of them are immunoreactive for parvalbumin (Jones 2007). In the cat, a quarter of the neurons in the MGV are GABAergic interneurons, but these are absent in rodents and bats (Winer and Larue 1988). The MGV neurons are well tuned to frequency, exhibiting low-threshold short-latency responses to tones and complex sounds. They



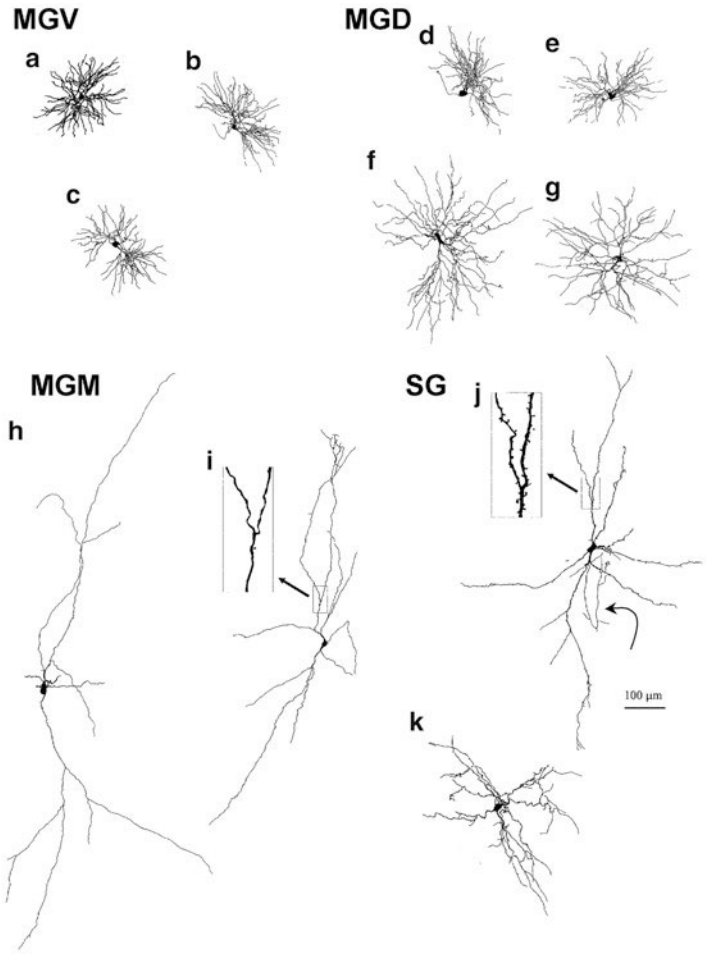
Anatomy and Physiology of the Mammalian Auditory System, Fig. 10 Schematic diagram of the basic circuitry in the IC. Large GABAergic neurons (*1*) receive excitatory inputs on their somata, send their axons to the MGB, and inhibit tufted or stellate neurons in the MGB. Small GABAergic neurons (*2*) do not project to MGB. Glutamatergic neurons (*3*, *4*) project to the MGB and lack the dense VGLUT2 axosomatic inputs. *Red* puncta indicate excitatory glutamatergic terminals. *Blue* puncta indicate inhibitory (GABAergic and glycinergic) terminals (Figure kindly provided by Dr. D. L. Oliver. Reproduced from Ito and Oliver 2012)



Anatomy and Physiology of the Mammalian Auditory System, Fig. 11 Distribution of calbindin and calretinin areas at two different rostrocaudal levels of the MGB in mouse. (**a** and **b**) Two sections from same mouse immunostained for calretinin, visualized with Alexa Fluor 594. (**c** and **d**) Identical sections from (**a** and **b**), immunostained for calbindin, visualized with Cy-2. (**e**) Overlay images from (**a** and **c**). (**f**) Overlay images from (**b** and **d**). Scale bar = 500 μm (Figure kindly provided by Dr. Daniel Llano. Adapted from Lu et al. 2009). Abbreviations in the figure: *PIL* Posterior intralaminar nucleus, *PP* Peripeduncular nucleus

are sensitive to both interaural time and intensity differences.

The MGV receives its ascending inputs primarily from the ipsilateral CNIC (Fig. 10), (Malmierca 2003; Jones 2007). These projections arise from both glutamatergic and GABAergic neurons (Bartlett and Smith 1999; Ito and Oliver 2012). MGV neurons project mainly to layers III



Anatomy and Physiology of the Mammalian Auditory System, Fig. 12 Camera lucida drawings of neurobiotin-labeled cells of the rat MG. Tufted cells populate the MGV. Cells with two distinct morphologies, tufted and stellate, populate MGD. Cells (**d** and **e**) and cells (**f** and **g**) are typical examples of MGD tufted and stellate cells, respectively. The dendritic trees of these cells appear similar regardless of the plane of section. Magnocellular neurons populate the MGM (Figure kindly provided by Dr. Philip Smith (some parts reproduced from Bartlett and Smith 1999)). Abbreviations in the figure: *SC* suprageniculate nucleus

and IV of the primary (core) areas of the auditory cortex.

The Dorsal Division of the Medial Geniculate Complex

The MGD is generally characterized by a lower cell-packing density than the MGV, and it lacks a laminar organization. At least two subdivisions are clearly recognized on the basis of architectonic variations and connections in most species (Burton and Jones 1976; Winer 1985). The most common neuronal type is the radiate cell (Fig. 12f–g). These have radially symmetric dendritic fields with a simple branching pattern. Tufted cells (Fig. 12a–c) are also present and tend to form thin sheets. Finally, there is a small population of small stellate cells. Compared to the MGV, parvalbumin immunoreactivity is reduced in the MGD, but there is a relative increase in the numbers of calbindin-positive cells (Jones 2007). GABAergic interneurons are also abundant in the MGD. Most MGD neurons respond over a wide range of latencies, typically longer than MGV neurons, and exhibit broader frequency tuning, so a clear tonotopic organization is not obvious, if it is present at all.

The main ascending inputs to the MGD arise from the DCIC and LCIC (Malmierca 2003; Jones 2007) and may be either glutamatergic or GABAergic in nature (Bartlett and Smith 1999). The MGD projects to the nonprimary (belt) areas of auditory cortex.

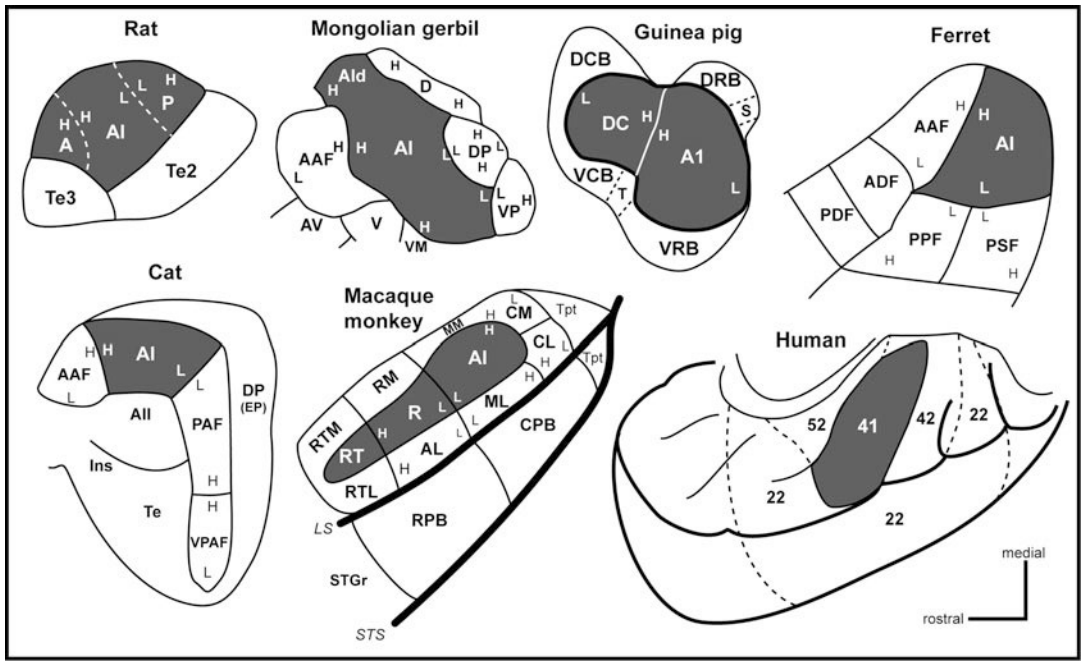
The Medial Division of the Medial Geniculate Complex

The MGM is located medial to the MGV and MGD and stretches from the rostral to caudal poles of the MGB (Fig. 11). The MGM is rather heterogeneous with respect to cell types and connections. Several different types of cells have been identified including the magnocellular type, the largest neurons in the entire MGB (Fig. 12h–i; Winer and Morest 1983). Some of the neurons are calbindin reactive and project mainly to layers I and II of cortex. Projections to the middle layers of cortex arise from both calbindin- and parvalbumin-positive neurons (Jones 2007). Some neurons are narrowly tuned to frequency and respond robustly at short latencies, similar to MGV neurons, while others are broadly tuned and have longer latencies. Some authors define a separate division referred to as suprageniculate nucleus (Fig. 12j–k).

MGM receives inputs from both auditory and nonauditory sources. The main auditory inputs arise from the cortical regions of the IC, as well as the CNC, SOC, and VLL (Anderson et al. 2006; Malmierca et al. 2002). Nonauditory inputs include the deep layers of the superior colliculus and other nuclei that appear to drive responses to somatic, vestibular, visual, and nociceptive stimuli in some species (see Jones 2007). In addition to auditory cortex, the MGM also projects to the striatum and amygdala (Doron and Ledoux 1999).

The Auditory Cortex

The auditory cortex (AC) is located in the temporal lobe of the cerebral hemisphere and represents the site of termination for fibers ascending from the auditory thalamus (Fig. 2). The most obvious species differences observed in the auditory brain are those seen in the AC (Fig. 13). For example, the number of areas identified in the AC ranges from 5 to 6 in mice and rats, 6–9 in cats and ferrets, 10–12 in primates, and over



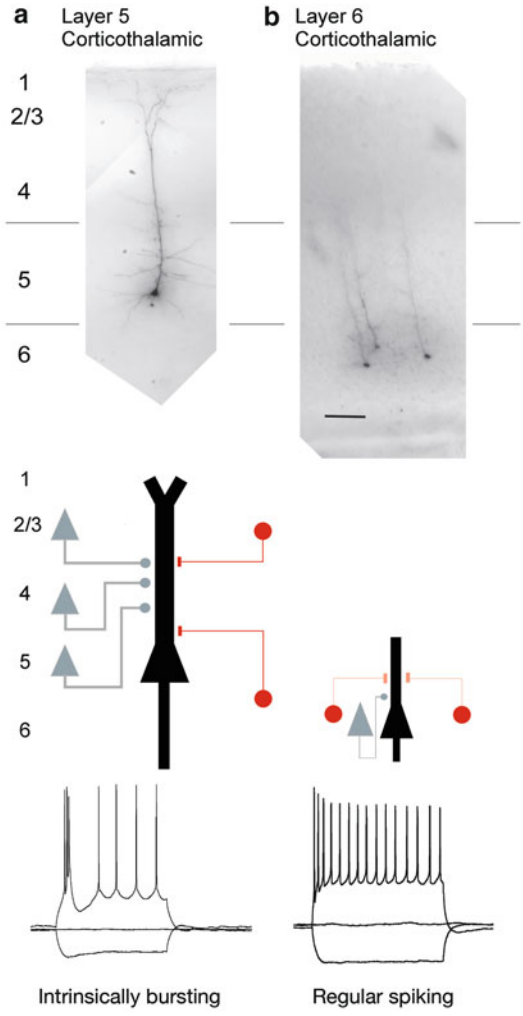
Anatomy and Physiology of the Mammalian Auditory System, Fig. 13 Schematics of auditory cortex models in several mammals. Primary (*core*) auditory areas are *shaded*. Tonotopic gradients are indicated by H (*high*) and L (*low*) frequency. (Reproduced from Malmierca and Hackett 2010)

30 in some studies of humans (Fig. 13). Species differences include the number of areas present, their relative position and arrangement, cell density, connections, and tonotopic organization, and it is likely that species differences will also be reflected in the organization of auditory cortex in other ways. However, a common theme is that a central primary region, or *core*, is surrounded by a variable number of secondary, or *belt*, areas in all mammals studied (Fig. 13). In nonhuman primates, the core–belt scheme has been extended to include a third region, known as the *parabelt* (Kaas and Hackett 1998; Winer and Schreiner 2011).

The Core Region of Auditory Cortex

The core region is made of the primary field (A1) and two other tonotopically organized areas (Fig. 13). Areas in the core are characterized by high cell density in a thick layer IV, dense myelination across laminae, and relatively high expression of several markers in the horizontal band involving layers III and IV, compared to secondary areas in the belt (Jones 2003; Kaas and Hackett 1998). Several different classes of pyramidal and nonpyramidal cells are found across the six layers of auditory cortex (Winer 1992). Pyramidal cells tend to be glutamatergic and are concentrated in layers III and V, while many nonpyramidal cells are GABAergic and account for about one-fourth of the neurons in most layers, except layer I, where they constitute more than 90%. Layer II contains both pyramidal and nonpyramidal neurons. The smaller cells are located superficially in layer II, while the larger pyramidal cells predominate near the border with layer III. Layer III is populated by several types of pyramidal and nonpyramidal cells. Layer IV is mainly populated by small tufted cells, which have radially oriented dendritic fields and are involved mainly in local columnar projections. Layer V contains both pyramidal and nonpyramidal neurons. The somata of the conspicuous large pyramidal neurons located in layer Vb have apical dendrites that extend to layer I, with several branches along the way. Other pyramidal cells in layer V are smaller and more evenly distributed. Layer V is of particular interest because its cells form part of the projection to the thalamus and other subcortical areas (Fig. 14; Games and Winer 1988; Hefti and Smith 2000; Malmierca and Ryugo 2011). As in other cortical regions, two distinct types of pyramidal cells, "intrinsically bursting" and "regular spiking" can be distinguished on the basis of their correlated morphology and physiology (Fig. 14; Kawaguchi 1993; Kasper et al. 1994). The *intrinsically bursting* pyramidal cells (Fig. 14) have large cell bodies and long, thick apical dendrites that branch extensively in layer I. Their axons arborize locally in the infragranular cortical layers and project into subcortical white matter. In contrast, the *regular-spiking* pyramidal cells (Fig. 14) have smaller cell bodies and a thinner apical dendrite that seldom extend to layer I. Their axons also project to the white matter and arborize locally in the supragranular cortex. In slice preparations, the intrinsically bursting neurons exhibit a characteristic firing pattern with a burst of action potentials followed by either additional bursts or single spikes, whereas the regular spiking neurons tend to fire single spikes with a variable degree of adaptation (Hefti and Smith 2000, 2003). Intrinsically bursting cells make up the majority of layer V's input to subcortical targets (Fig. 15; Games and Winer 1988; Winer 1992; Weedman and Ryugo 1996a, b; Saldaña et al. 1996) and are capable of providing a robust input to postsynaptic neurons (Hefti and Smith 2000). In contrast, most regular spiking neurons are strongly inhibited and may provide less robust but perhaps more specific, information to their inputs. Finally, layer VI has the widest variety of cell types, including several classes of pyramidal cells and multipolar, bipolar, and horizontal cells.

The main source of ascending inputs to A1 and other core areas is the MGV (Fig. 2; Jones 2007; Winer 1985; Winer and Lee 2007). The thalamocortical termination is concentrated in layers III and IV. The organization of these connections is topographic, reflecting tonotopic organization within the MGV, as well as the areas to which it projects. Additional inputs to the core



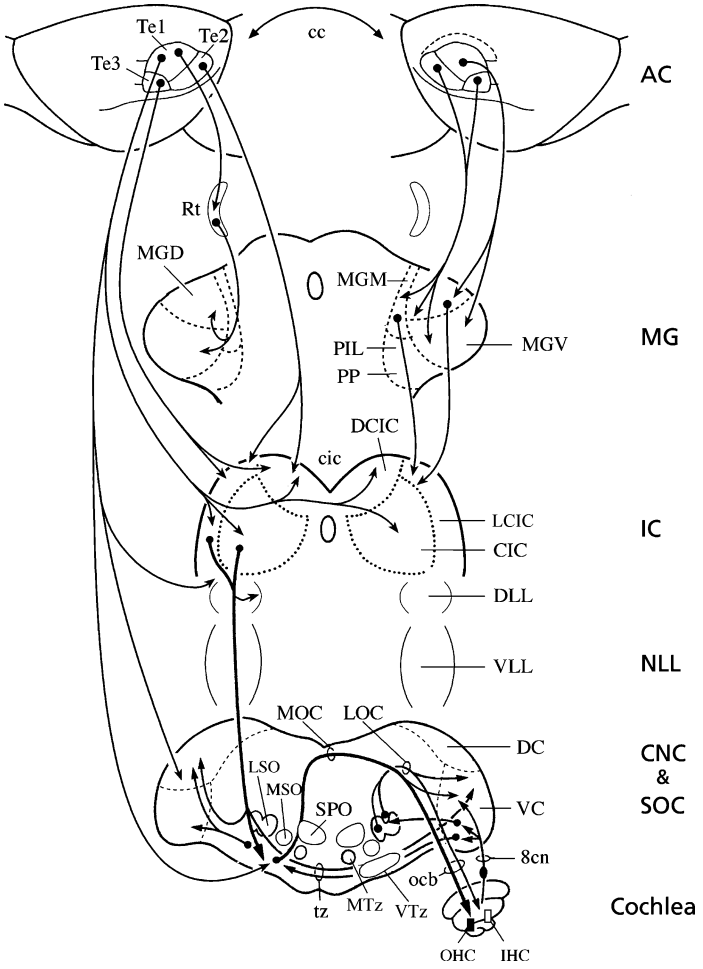
Anatomy and Physiology of the Mammalian Auditory System, Fig. 14 *Top panel*, examples of morphology of layer V and layer VI corticothalamic neurons biocytin-filled corticothalamic neurons. Scale bar 100 μm; *Bottom panels*, Llano and Sherman model of differences of synaptic input and response properties between auditory layer V and layer VI corticothalamic neurons. In response to excitatory input, layer V corticothalamic neurons fire a burst of action potentials at low threshold, and these neurons receive excitatory input from neurons in layers II/III, IV, and V (*gray*) and GABAA-mediated inhibitory input mostly from lower layer V with a smaller contribution from layer II/III (*red*). In contrast, in response to excitatory input, these neurons fire a regular train of individual action potentials, and these neurons receive excitatory input primarily from layer VI and inhibitory input from adjacent areas in layer VI (Figure kindly provided by Dr. Daniel Llano. Adapted from Llano and Sherman 2009)

include MGM, which projects broadly to all areas of auditory cortex. The intracortical connections of the core mainly include other areas of auditory cortex ipsilaterally and sparse connections with areas beyond auditory cortex. Tonotopically matched sites are more densely interconnected than non-matched sites (Lee and Winer 2005). Connections with auditory cortex in the opposite hemisphere are concentrated in the homotopic (matching) area, (Wallace and Harper 1997) and heterotopic connections are relatively weak. The main callosal projections arise from both pyramidal and nonpyramidal cells in layers III and V. Layers V and VI represent the main source of descending projections to the MGB and IC and brainstem (Fig. 15).

The Belt Region of Auditory Cortex

Areas that lie outside of the core region are often referred to as the nonlemniscal or belt areas (Fig. 13). These areas are anatomically and physiologically distinct from the core fields and from one another. Architectonically, each of the belt areas is distinct, but compared to the core region, cell density and myelination are generally reduced, as is the expression of cytochrome oxidase, acetylcholinesterase, and parvalbumin (Jones 2003).

In addition to the inputs from the core, the main source of projections to most of the belt areas is the MGD (Jones 2007; Winer 1985; Winer and Lee 2007). Additional connections include the MGM and thalamic nuclei that surround the MGB. Belt areas tend to differ with respect to the balance of inputs from different thalamic nuclei. In nonhuman primates, an additional group of areas has been identified that surround the belt areas adjacent to the core. This region is known as the *parabelt* (Kaas and Hackett 1998). The parabelt region receives inputs from the belt region and MGD, but not the core. Thus, some ascending information appears to pass serially through the core and belt regions before reaching the parabelt (Kaas and Hackett 1998).

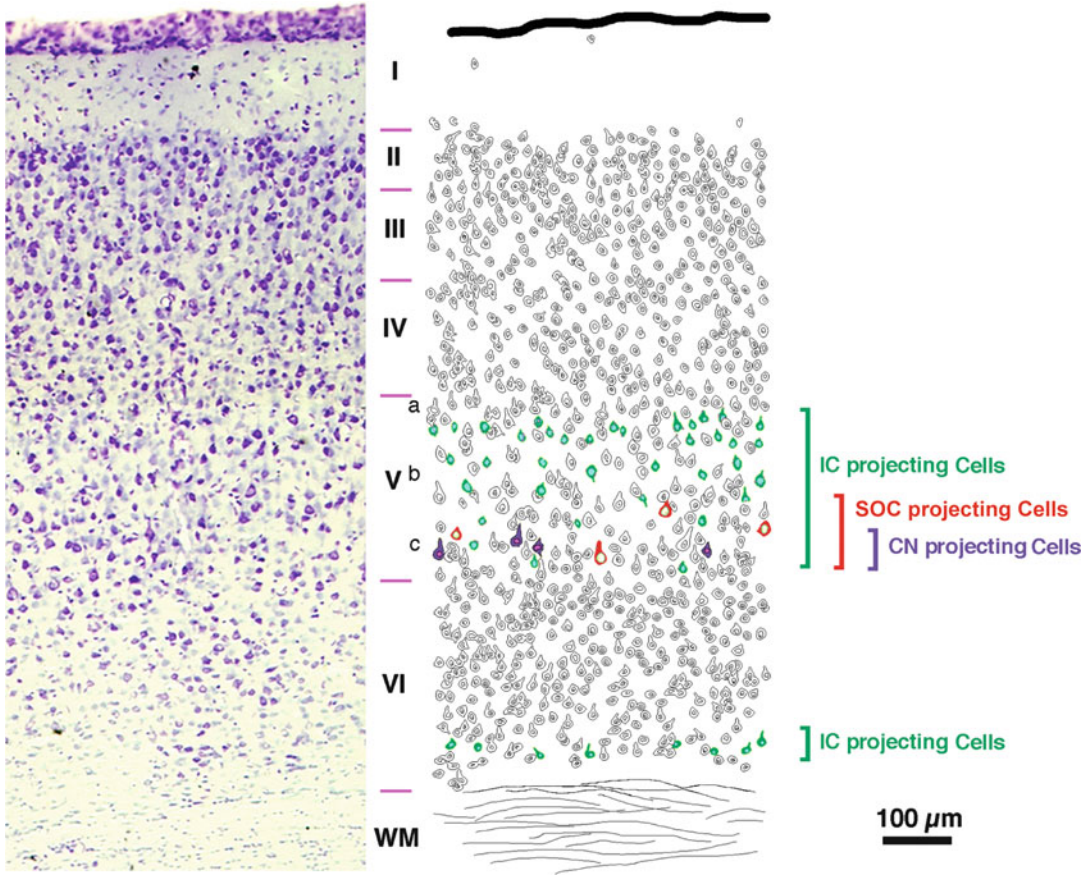


Anatomy and Physiology of the Mammalian Auditory System, Fig. 15 Schematic wiring diagram of the descending auditory pathway of the rat (Modified after Brodal 1981, AC is from Herbert et al. 1991). Abbreviations in the figure: *bic* Brachium of the inferior colliculus, *cc* Corpus callosum, *CIC* Central nucleus of the inferior colliculus, *cic* Commissure of the inferior colliculus, *cll* Commissure of the lateral lemniscus (Prosbt), *das* Dorsal acoustic stria, *DC* Dorsal cochlear nucleus, *h* High-frequency region, *IC* Inferior colliculus, *l* Low-frequency region, *ll* Lateral lemniscus, *LTz* Lateral nucleus of the trapezoid body, *MG* Medial geniculate body, *MTz* Medial nucleus of the trapezoid body, *ocb* Olivocochlear bundle, *PIL* Posterior intralaminar nucleus, *PP* Peripeduncular nucleus, *Rt* Auditory sector of the reticular thalamic nucleus, *SOC* superior olivary nucleus, *SPO* Superior paraolivary nucleus, *Te1* Temporal area 1, *Te2* Temporal area 2, *Te3* Temporal area 3, *tz* Trapezoid body (or ventral acoustic stria), *VC* Ventral cochlear nucleus, *8cn* Cochlear root of the vestibulocochlear nerve

The Descending Auditory Pathways

In parallel to the ascending auditory pathways, there are stepwise, descending projections from the auditory cortex to the organ of Corti (Fig. 15). The descending auditory pathway could be considered to consist of both (1) a descending chain of connections and (2) a series of regional feedback loops (Warr 1992).

The descending chain comprises three main levels. The first level originates in the AC and includes corticothalamic, corticotectal, and corticopontine projections (Fig. 16). The corticothalamic projection forms reciprocal and



Anatomy and Physiology of the Mammalian Auditory System, Fig. 16 Diagrammatic summary of the laminar organization of cortical cells projecting to the inferior colliculus (*IC*), superior olivary complex (*SOC*), and cochlear nucleus (*CN*). All three distributions overlap; however, the cortical neurons projecting to more distant targets are more narrowly distributed and centered in deeper regions of layer V (Figure kindly provided by Dr. D. K. Ryugo. Reproduced from Doucet et al. 2003)

nonreciprocal connections between the AC and MGB. The corticotectal projection terminates in the IC and subcollicular nuclei (Feliciano and Potashner 1995; Meltzer and Ryugo 2006). The second level of the chain originates in the IC, whose descending fibers form colliculo-olivary and colliculo-cochleonuclear projections (Fig. 15). The colliculo-olivary fibers terminate on the periolivary medial olivocochlear cells which supply the OHCs (Caicedo and Herbert 1993; Faye-Lund 1985; Vetter et al. 1993). They may also terminate on lateral olivocochlear cells, which supply the IHCs (Feliciano and Potashner 1995). The last and third level of the chain consists of the olivocochlear system that provides efferent innervation to the cochlea (Figs. 15 and 19).

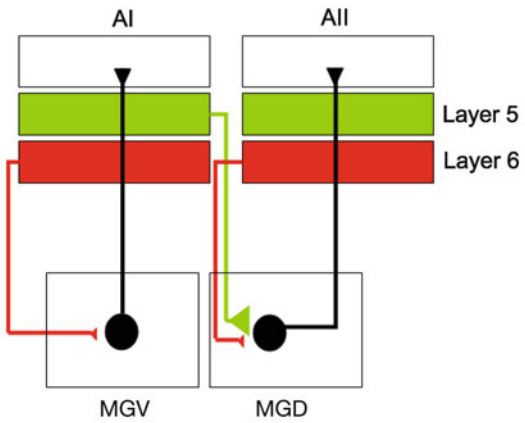
The regional feedback loops consist of a series of cortical projections to subcortical nuclei that project back to cortex, directly or indirectly, allowing AC to control its inputs from lower centers.

The Corticofugal Pathways

The major AC projections target the MGB and IC (Figs. 15–18), but there are also AC projections to subcollicular nuclei, including the nucleus sagulum, paralemniscal regions, superior olivary complex, cochlear nuclear complex, and pontine

nuclei. The AC also supplies the amygdala, basal ganglia, striatum, superior colliculus, and central gray (reviewed in Malmierca and Ryugo 2011, 2012), suggesting that the AC has important roles in addition to auditory sensory processing (Winer 2006).

In the auditory thalamus, the MGV receives its heaviest source of input from the primary AC, while the MGM the least and the MGD receives an intermediate amount. A prominent feature of the corticothalamic projection is reciprocity, in which a cortical region projects to the part of the thalamus from which it receives input. However, there are also nonoverlapping regions (Llano and Sherman 2008). In the MGB, there are two main types of terminal synaptic boutons (I and II) arising from the core of the AC (Fig. 17; Bartlett et al. 2000; Llano and Sherman 2009). Type I terminals are small (<1 μm in long-axis diameter), synapse on small caliber dendrites of the MGV and MGD, and arise from the pyramidal cells of layer VI. Type II terminals are large (>2 μm in long-axis diameter), mostly synapse in the MGD and occasionally in the MGM (Bartlett et al. 2000; Llano and Sherman 2009) and arise from pyramidal neurons from layer V.



Anatomy and Physiology of the Mammalian Auditory System, Fig. 17 Diagram of a corticothalamocortical model of cortical processing. Information reaches the *lemniscal* AC via a projection from the MGV to layer IV of either the AAF or the AI. From here, a layer 5 pyramidal projects to the MGD, where a thalamocortical relay cell projects to layer IV of the *nonlemniscal* AC (DP or AII). From the nonlemniscal AC, a layer VI projection is sent to the MGB, adhering to the principle that all thalamocortical projections, whether coming from the first- or higher-order thalamus, receive a modulator, reciprocal projection from layer VI (Figure kindly provided by Dr. Daniel Llano. Adapted from Llano and Sherman 2008)

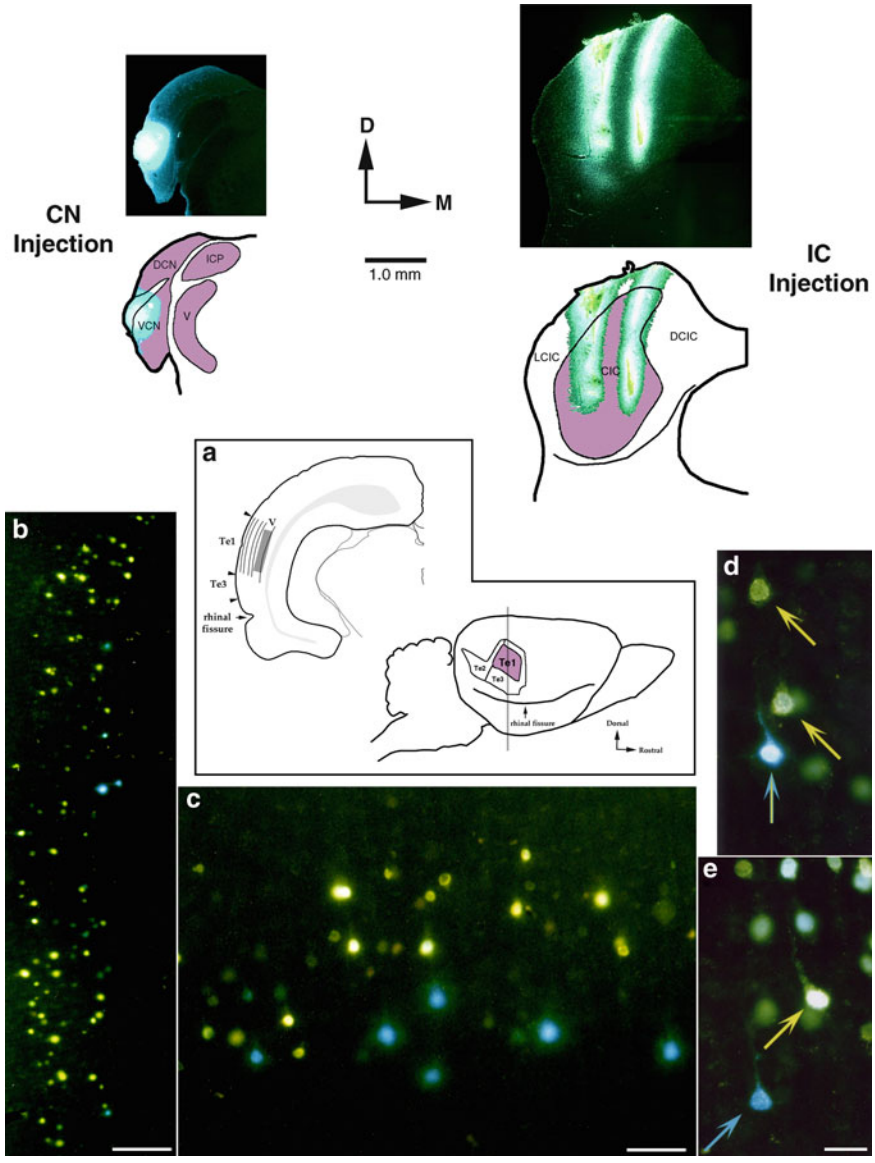
Sherman and Guillery (1996) first proposed the notion of "drivers" and "modulators" of thalamic neurons in the visual and somatosensory thalamus, but this hypothesis has been applied to the auditory system as well (Llano and Sherman 2008). According to this theory, type I terminals play a modulatory role in the first-order thalamic nuclei, such as the MGV (Fig. 17). Thus, the corticothalamic inputs converge with ascending inputs on thalamic neurons such that the ascending inputs drive the thalamic neurons and the cortical inputs modulate them. In contrast, in the "higher-order" thalamic nuclei, such as the MGD, the "driver" inputs arise from the large type II axons and terminals that originate from the cortex and interact with other ascending input from the IC.

The corticofugal projection is glutamatergic and modulates the MGB responses to sound through a direct excitatory pathway, but the AC can also provide the MGB with an inhibitory influence (Bartlett et al. 2000) via its projections to the auditory sector of the thalamic reticular nucleus which, in turn, projects to the MGB.

As in the MGB, AC may modulate the processing of sounds in the IC through the activation of local inhibitory connections within the IC. Several studies have shown direct neocortical projections to regions surrounding the lateral lemniscus, including the nucleus sagulum ipsilaterally and the SOC and CNC bilaterally (Feliciano and Potashner 1995; Weedman et al. 1996; Weedman and Ryugo 1996a, b).

The Colliculofugal Pathways

Studies in several species have shown that the IC has descending projections to the LL, SOC, and CNC (Fig. 15; reviewed in Malmierca and Ryugo 2011, 2012). Perhaps, the most interesting projections are the *colliculo-olivary* projections that originate in the CNIC and LCIC because they terminate on the VNTB, the site of origin of the



Anatomy and Physiology of the Mammalian Auditory System, Fig. 18 Retrograde labeling in area Te1 for a rat that received an injection of Fast *blue* (FB) into the CNC and Diamidino *yellow* (DiY) into the IC (*top*). (**a**) Sagittal view of the brain, the *gray line* through AC indicates the position of the cells along the rostral/caudal axis display of the location of labeled cortical cells shown in panels (**b**–**e**). (**b**) Photomontage of layer V with few FB-labeled cells located in deep layer V whereas the DiY-labeled neurons are distributed more broadly. (**c**) Higher magnification photomontage illustrating the laminar organization of cortical cells (cortical surface is towards the *top*) projecting to the CNC (*blue*) vs those targeting the IC (*yellow*) within layer V. (**d** and **e**) Examples of labeled cortical cells. Cortical cells that contained both dyes (*blue and yellow arrow* in **d**) were observed much less frequently (Figure kindly provided by Dr. D. K.Ryugo). Reproduced from Doucet et al. 2003). Abbreviations in the figure: *CN* Cochlear nucleus, *CIC* Central nucleus of the inferior colliculus, *D* Dorsal, *M* Medial, *Te1* Temporal area 1, *Te2* Temporal area 2, *Te3* Temporal area 3

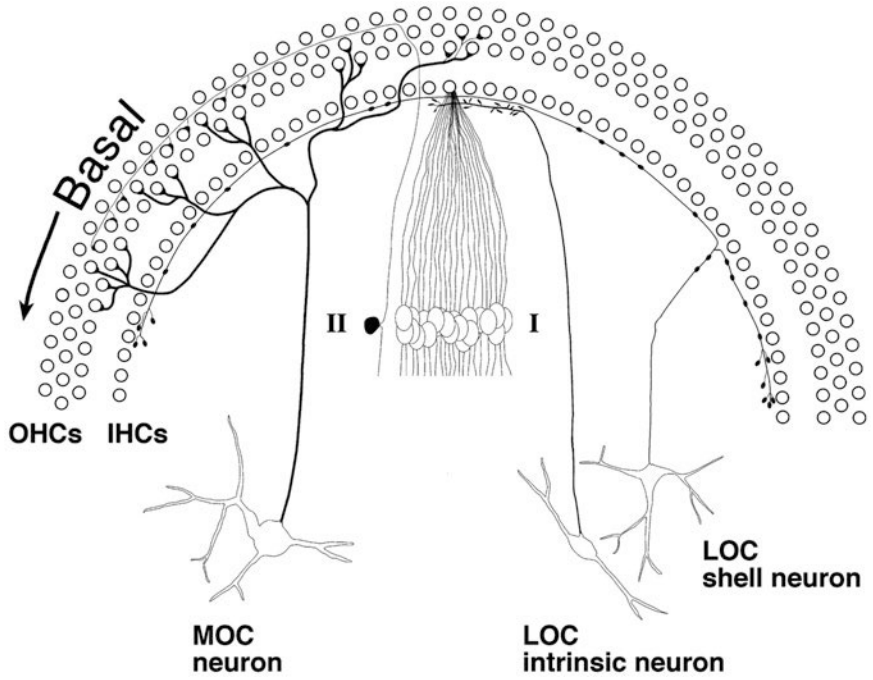
MOC (White and Warr 1983), suggesting that the IC influences MOC neurons (Vetter et al. 1993). The IC also projects to nonauditory nuclei including the pontine nuclei, lateral paragigantocellular nucleus, gigantocellular reticular nucleus, ventrolateral tegmental nucleus, and caudal pontine reticular nucleus (Caicedo and Herbert 1993).

The Olivocochlear System

The olivocochlear bundle provides the organ of Corti with efferent innervation (Figs. 15 and 19; Rasmussen 1946) and plays a critical role in maintaining the normal operation of the cochlea (Figs. 15 and 19). It may introduce nonlinear dynamics into the auditory system (Eggermont 2001). There are two systems of olivocochlear neurons, medial (MOC) and lateral (LOC). The MOC neurons are located medial and ventral to the main nuclei of the SOC in the PO nucleus known as the VNTB and project mainly to the contralateral cochlea, whereas the LOC neurons are located within or near the LSO and project to the ipsilateral cochlea (Warr 1992; White and Warr 1983).

The MOC neurons constitute a homogeneous population of cholinergic cells and innervate the OHCs (Figs. 15 and 19). About 75% of them originate on the contralateral side with the remainder originating ipsilaterally (Warr 1992; Brown and Levine 2008). They probably receive descending input from the ipsilateral IC and ascending input bilaterally from the VCN.

In rodents, the LOC neurons (Figs. 15 and 19) consist of two distinct types of neurons: *intrinsic* and *shell* neurons (Vetter and Mugnaini 1992; Warr et al. 1997). *Intrinsic neurons* are confined to the ipsilateral LSO and constitute about 85% of all LOC neurons (Brown and Levine 2008; Vetter and Mugnaini 1992; Warr et al. 1997). Most *shell neurons* surround the ipsilateral LSO and constitute the remaining 15% (Vetter and Mugnaini 1992; Brown and Levine 2008; Warr et al.



Anatomy and Physiology of the Mammalian Auditory System, Fig. 19 Scheme of the LOC and MOC neurons, their projections to the IHCs and OHCs, respectively, and afferent fiber types I and II projecting to the CNC. In rodents, two types of LOC cells occur: intrinsic located inside the LSO and shell located in the margins of the LSO (Figure kindly provided by Dr. Bruce Warr. Adapted from Warr 1992)

1997). (In contrast, in the cat the LOC neurons surround the LSO, and two populations have not been distinguished.) The LOC neurons innervate the type I primary afferent fibers near the region where they contact the IHCs. Virtually all of them originate on the ipsilateral side (Brown and Levine 2008).

The functional role of the MOC neurons may be to enhance transduction or signal detection through an unmasking effect, thus regulating the slow motility of the OHCs and thereby the stiffness of the basilar membrane (Eggermont 2001). Another function may be to protect the inner ear from acoustic injury (Taranda et al. 2009). Although the functional role of the LOC is uncertain, several studies (Safieddine et al. 1997; Safieddine and Eybalin 1992) have shown that these neurons may provide a modulatory effect to the afferent fibers that contact the IHCs.

Cross-References

- ▶ Acoustic Timbre Recognition
- ▶ Associations and Rewards in the Auditory Cortex
- ▶ Attentional Top-Down Modulation, Models of
- ▶ Auditory Event-Related Potentials
- ▶ Auditory Evoked Brainstem Responses
- ▶ Auditory Memory
- ▶ Auditory Perceptual Organization
- ▶ Auditory Precedence Effect
- ▶ Auditory Processing in Insects
- ▶ Auditory Prosthesis
- ▶ Auditory Sensing Systems: Overview
- ▶ Auditory Thalamocortical Transformations
- ▶ Auditory-Nerve Response, Afferent Signals
- ▶ Brainstem Processing: Overview
- ▶ Cochlear Inner Hair Cell, Model
- ▶ Context-Dependent Processing in Auditory Cortex
- ▶ Cortical Maps, Activity-Dependent Development
- ▶ Local Field Potentials: LFP
- ▶ Music Processing in the Brain
- ▶ Neural Coding of Speech Sounds
- ▶ Neuromorphic Sensors, Cochlea
- ▶ Peripheral Nerve Interface Applications, Cochlear Implants
- ▶ Pitch Perception, Models
- ▶ Receptive Field Modeling
- ▶ Sensory Coding, Efficiency
- ▶ Sound Localization and Experience-Dependent Plasticity
- ▶ Sound Localization in Mammals and Models
- ▶ Spectrotemporal Receptive Fields
- ▶ Spike-Frequency Adaptation
- ▶ Stimulus-Specific Adaptation, Models
- ▶ Stimulus-Specific Information
- ▶ Tinnitus, Models

Acknowledgments I am most thankful and indebted to Dr. Nell Cant for her suggestions and constructive criticisms on a previous version of the manuscript.
Financial support was provided by Spanish MINECO (BFU2009-07286) and EU (EUI2009-04083, in the framework of the ERA-NET Network of European Funding for Neuroscience Research).

References

Adams JC (1979) Ascending projections to the inferior colliculus. J Comp Neurol 183:519–538
Adams JC (1983) Cytology of periolivary cells and the organization of their projections in the cat. J Comp Neurol 215:2752–2789
Aitkin LM, Dickhaus H, Schult W, Zimmermann M (1978) External nucleus of inferior colliculus, auditory and spinal somatosensory afferents and their interactions. J Neurophysiol 41:837–847
Anderson LA, Malmierca MS, Wallace MN, Palmer AR (2006) Evidence for a direct, short latency projection from the dorsal cochlear nucleus to the auditory thalamus in the Guinea pig. Eur J Neurosci 24:491–498
Banks MI, Smith PH (1992) Intracellular recordings from neurobiotin-labeled cells in brain slices of the rat medial nucleus of the trapezoid body. J Neurosci 12: 2819–2837
Bartlett EL, Smith PH (1999) Anatomic, intrinsic, and synaptic properties of dorsal and ventral division neurons in rat medial geniculate body. J Neurophysiol 81: 1999–2016
Bartlett EL, Stark JM, Guillery RW, Smith PH (2000) Comparison of the fine structure of cortical and collicular terminals in the rat medial geniculate body. Neuroscience 100:811–828
Blaesse P, Ehrhardt S, Friauf E, Nothwang HG (2005) Developmental pattern of three vesicular glutamate transporters in the rat superior olivary complex. Cell Tissue Res 320:33–50

Brawer JR, Morest DK, Kane EC (1974) The neuronal architecture of the cochlear nucleus of the cat. J Comp Neurol 155:251–300

Brodal A (1981) Neurological anatomy in relation to clinical medicine, 3rd edn. Oxford University Press, Oxford

Brown MC, Levine JL (2008) Dendrites of medial olivocochlear neurons in mouse. Neuroscience 154: 147–159

Brown MC, Liberman MC, Benson TE, Ryugo DK (1988) Brainstem branches from olivocochlear axons in cats and rodents. J Comp Neurol 278:591–603

Burton H, Jones EG (1976) The posterior thalamic region and its cortical projection in New World and Old World monkeys. J Comp Neurol 168:249–301

Caicedo A, Herbert H (1993) Topography of descending projections from the inferior colliculus to auditory brainstem nuclei in the rat. J Comp Neurol 328: 377–392

Cant NB, Benson CG (2003) Parallel auditory pathways: projection patterns of the different neuronal populations in the dorsal and ventral cochlear nuclei. Brain Res Bull 60:457–474

Cant NB, Benson CG (2006) Organization of the inferior colliculus of the gerbil (*Meriones unguiculatus*): differences in distribution of projections from the cochlear nuclei and the superior olivary complex. J Comp Neurol 495:511–528

Cant NB, Benson CG (2007) Multiple topographically organized projections connect the central nucleus of the inferior colliculus to the ventral division of the medial geniculate nucleus in the gerbil, *Meriones unguiculatus*. J Comp Neurol 503:432–453

Casseday JH, Fremouw T, Covey E (2002) The inferior colliculus, a hub for the central audiotory system. In: Oertel D, Popper AN, Fay RR (eds) Springer handbook of auditory research. Springer, New York

Cotillon N, Nafati M, Edeline JM (1999) Characteristics of reliable tone-evoke oscillations in the rat thalamocortical auditory system. Hear Res 142:113–130

Covey E, Casseday JH (1991) The monaural nuclei of the lateral lemniscus in an echolocating bat, parallel pathways for analyzing temporal features of sound. J Neurosci 11:3456–3470

Dehmel S, Kopp-Scheinpflug C, Dörrscheidt GJ, Rübsamen R (2002) Electrophysiological characterization of the superior paraolivary nucleus in the *Mongolian gerbil*. Hear Res 172:18–36

Doron NN, Ledoux JE (1999) Organization of projections to the lateral amygdala from auditory and visual areas of the thalamus in the rat. J Comp Neurol 412:383–409

Doucet JR, Ryugo DK (1997) Projections from the ventral cochlear nucleus to the dorsal cochlear nucleus in rats. J Comp Neurol 385:245–264

Doucet JR, Ross AT, Gillespie MB, Ryugo DK (1999) Glycine immunoreactivity of multipolar neurons in the ventral cochlear nucleus which project to the dorsal cochlear nucleus. J Comp Neurol 408:515–531

Doucet JR, Molavi DL, Ryugo DK (2003) The source of corticocollicular and corticobulbar projections in area Te1 of the rat. Exp Brain Res 4:461–466

Echteler SM, Fay RR, Popper AN (1994) Structure of the mammalian cochlea. In: Fay RR, Popper AN (eds) Comparative hearing, mammals. Springer, Berlin, pp 134–171

Eggermont JJ (2001) Between sound and perception, reviewing the search for a neural code. Hear Res 157: 1–42

Fay RR (1988) Hearing in vertebrates: a psychophysics databook. Hill-Fay Associates, Winnetka

Fay RR, Popper AN (1994) Comparative hearing in mammals. In: Fay RR, Popper AN (eds) Springer handbook of auditory research. Springer, New York

Faye-Lund H (1985) The neocortical projection to the inferior colliculus in the albino rat. Anat Embryol (Berl) 173:53–70

Faye-Lund H, Osen KK (1985) Anatomic of the inferior colliculus in rat. Anat Embryol 175:35–52

Feliciano M, Potashner SJ (1995) Evidence for a glutamatergic pathway from the Guinea pig auditory cortex to the inferior colliculus. J Neurochem 65: 1348–1357

Friauf E (1993) Transient appearance of calbindin-D28k-positive neurons in the superior olivary complex of developing rats. J Comp Neurol 334:59–74

Games KD, Winer JA (1988) Layer V in rat auditory cortex, projections to the inferior colliculus and contralateral cortex. Hear Res 34:1–25

Guinan JJ, Norris BE, Guinan SS (1972) Single auditory units in the superior olivary complex, II. Locations of unit categories and tonotopic organization. Int J Neurosci 4:147–166

Hefti BJ, Smith PH (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

Hefti BJ, Smith PH (2003) Distribution and kinetic properties of GABAergic inputs to layer V pyramidal cells in rat auditory cortex. J Assoc Res Otolaryngol 4: 106–121

Held H (1893) Die centralem Bahnen des Nervus acusticus bei der. Katz Arch Anat Abtheil 15:190–271

Helfert RH, Aschoff A (1997) Superior olivary complex and nuclei of the lateral lemniscus. In: Ehret G, Romand R (eds) Anatomical and functional aspects of the cochlear nucleus. Oxford University Press, Oxford, pp 193–257

Herbert H, Aschoff A, Ostwald J (1991) Topography of projections from the auditory cortex to the inferior colliculus in the rat. J Comp Neurol 304:103–122

Irvine DRF (1992) Physiology of the auditory brainstem. In: Popper AN, Fay RR (eds) Springer handbook of auditory pathway, neurophysiology. Springer, New York, pp 153–231

Ito T, Oliver DL (2012) Patterns of synaptic organization send different messages to the thalamus. Front Neural Circuits 6:48. https://doi.org/10.3389/fncir.2012.00048

Jones EG (2003) Chemically defined parallel pathways in the monkey auditory system. Ann N Y Acad Sci 999: 218–233

Jones EG (2007) The thalamus, vol II, 2nd edn. Cambridge University Press, Cambridge, pp 875–923

Joris PX, Smith PH, Yin TCT (1998) Coincidence detection in the auditory system, 50 years after Jeffress. Neuron 21:1235–1238

Kaas JH, Hackett TA (1998) Subdivisions of auditory cortex and levels of processing in primates. Audiol Neurootol 3:73–85

Kasper EM, Larkman AU, Lubke J, 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

Kawaguchi Y (1993) Groupings of nonpyramidal and pyramidal cells with specific physiological and morphological characteristics in rat frontal cortex. J Neurophysiol 69:416–431

Kelly JB, Caspary DM (2005) Pharmacology of the inferior colliculus. In: Winer JA, Schreiner C (eds) The inferior colliculus. Springer, New York, pp 248–281

King AJ, Jiang ZD, Moore DR (1998) Auditory brainstem projections to the ferret superior colliculus: anatomical contribution to the neural coding of sound azimuth. J Comp Neurol 390:342–365

Kopp-Scheinpflug C, Tolnai S, Malmierca MS, Rübsamen R (2008) The medial nucleus of the trapezoid body: comparative physiology. Neuroscience 154:60–170

Kulesza RJ Jr, Spirou GA, Berrebi AS (2003) Physiological response properties of neurons in the superior paraolivary nucleus of the rat. J Neurophysiol 89: 2299–2312

LeBeau FEN, Malmierca MS, Rees A (2001) Iontophoresis in vivo demonstrates a key role for $GABA_A$- and glycinergic inhibition in shaping frequency response areas in the inferior colliculus of Guinea pig. J Neurosci 21:7303–7312

Lee CC, Winer JA (2005) Principles governing auditory cortex connections. Cereb Cortex 15:1804–1814

Liberman MC, Dodds LW, Pierce S (1990) Afferent and efferent innervation of the cat cochlea, quantitative analysis with light and electron microscopy. J Comp Neurol 301:443–460

Llano DA, Sherman SM (2008) Evidence for nonreciprocal organization of the mouse auditory thalamocortical-corticothalamic projection systems. J Comp Neurol 507:1209–1227

Llano DA, Sherman SM (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

Loftus B, Malmierca MS, Oliver DL (2008) The Cytoarchitecture of the inferior colliculus revisited: a common organization of the lateral cortex in rat and cat. Neuroscience 154:196–205

Lorente de Nó R (1981) The primary acoustic nuclei. Raven Press, New York

Lu E, Llano DA, Sherman SM (2009) Different distributions of calbindin and calretinin immunostaining across the medial and dorsal divisions of the mouse medial geniculate body. Hear Res 257:16–23

Malmierca MS (2003) The structure and physiology of the rat auditory system: an overview. Int Rev Neurobiol 56: 147–211

Malmierca MS, Hackett TA (2010) Structural organization of the ascending auditory pathway. In: Moore DR (ed) The Oxford handbook of auditory science: the auditory brain. Oxford University Press, New York, pp 9–41

Malmierca MS, Ryugo DK (2011) Descending connections of auditory cortex to the midbrain and brainstem. In: Winer JA, Schreiner CE (eds) The auditory cortex. Springer, New York, pp 189–208

Malmierca MS, Ryugo DK (2012) Auditory system. In: Watson C, Paxinos G, Puelles L (eds) The mouse nervous system. Academic, Amsterdam, pp 607–645

Malmierca MS, Blackstad TW, Osen KK, Karagülle T, Molowny RL (1993) The central nucleus of the inferior colliculus in rat, a Golgi and computer reconstruction study of neuronal and laminar structure. J Comp Neurol 333:1–27

Malmierca MS, Rees A, LeBeau FEN, Bajaalie JG (1995) Laminar organization of frequency-defined local axons within and between the inferior colliculi of the Guinea pig. J Comp Neurol 357:124–144

Malmierca MS, LeBeau FEN, 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

Malmierca MS, Leergard TB, Bajo VM, Bjaalie JG (1998) Anatomic evidence of a 3-D mosaic pattern of tonotopic organization in the ventral complex of the lateral lemniscus in cat. J Neurosci 19:10603–10618

Malmierca MS, Merchán M, Henkel CK, Oliver DL (2002) Direct projections from the dorsal cochlear nucleus to the auditory thalamus in rat. J Neurosci 22: 10891–10897

Malmierca MS, Saint Marie RL, Merchan MA, Oliver DL (2005) Laminar inputs from dorsal cochlear nucleus and ventral cochlear nucleus to the central nucleus of the inferior colliculus: two patterns of convergence. Neuroscience 136:883–894

Malmierca MS, Izquierdo MA, Cristaudo S, Hernández O, Pérez-González D, Covey E, Oliver DL (2008) A discontinuous tonotopic organization in the inferior colliculus of the rat. J Neurosci 28:4767–4776

Malmierca MS, Cristaudo S, Pérez-González D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. J Neurosci 29: 5483–5493

Meltzer NE, Ryugo DK (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

Merchán M, Aguilar LA, Lopez-Poveda EA, Malmierca MS (2005) The inferior colliculus of the rat: quantitative immunocytochemical study of GABA and glycine. Neuroscience 136:907–925

Moore JK, Osen KK (1979) The human cochlear nuclei. In: Creutzfeld O, Scheich H, Schreiner C (eds) Experimental Brain Research, Suppementum II, pp 36–44

Morest DK (1968) The collateral system of the medial nucleus of the trapezoid body of the cat, its neuronal architecture and relation to the olivo-cochlear bundle. Brain Res 9:288–311

Mugnaini E, Osen KK, Dahl A-L, Friedrich VL Jr, Korte G (1980a) Fine structure of granule cells and related interneurons (termed Golgi cells) in the cochlear nuclear complex of cat, rat and mouse. J Neurocytol 9:537–570

Mugnaini E, Warr WB, Osen KK (1980b) Distribution and light microscopic features of granule cells in the cochlear nuclei of cat, rat, and mouse. J Comp Neurol 191:581–606

Oertel D (1999) The role of timing in the brain stem auditory nuclei of vertebrates. Annu Rev Physiol 61: 497–519

Oertel D, Young ED (2004) What's a cerebellar circuit doing in the auditory system? Trends Neurosci 27: 104–110

Oertel D, Wu SH, Garb MW, Dizack C (1990) Morphology and physiology of cells in slice preparations of the posteroventral cochlear nucleus of mice. J Comp Neurol 295:136–154

Oliver DL, Morest DK (1984) The central nucleus of the inferior colliculus in the cat. J Comp Neurol 222: 237–264

Oliver DL, Ostapoff EM, Beckius GE (1999) Direct innervation of identified tectothalamic neurons in the inferior colliculus by axons from the cochlear nucleus. Neuroscience 93:643–658

Osen KK (1969) Cytoarchitecture of the cochlear nuclei in the cat. J Comp Neurol 136:453–483

Osen KK, Mugnaini E, Dahl AL, Christiansen AH (1984) Histochemical localization of acetylcholinesterase in the cochlear and superior olivary nuclei. A reappraisal with emphasis on the cochlear granule cell system. Arch Ital Biol 122:169–212

Osen KK, Ottersen OP, Størm-Mathisen J (1990) Colocalization of glicine-like and GABA-like inmunoreactivities, a semiquantitative study of individual neurons in the dorsal cochlear nucleus of cat. In: Ottersen OP, Størm-Mathissen J (eds) Glycine neurotransmission. Wiley, Chichester, pp 417–451

Ota Y, Oliver DL, Dolan DF (2004) Frequency-specific effects on cochlear responses during activation of the inferior colliculus in the Guinea pig. J Neurophysiol 91: 2185–2193

Palombi PS, Caspary DM (1996) Responses of young and aged Fischer 344 rat inferior colliculus neurons to binaural tonal stimuli. Hear Res 100:59–67

Peruzzi D, Bartlett E, Smith PH, Oliver DL (1997) A monosynaptic GABAergic input from the inferior colliculus to the medial geniculate body in rat. J Neurosci 17:3766–3777

Rajan R (1990) Electrical stimulation of the inferior colliculus at low rates protects the cochlea from auditory desensitization. Brain Res 506: 192–204

Rasmussen GL (1946) The olivary peduncle and other fiber projections of the superior olivary complex. J Comp Neurol 99:61–74

Rietzel HJ, Friauf E (1998) Neuron types in the rat lateral superior olive and developmental changes in the complexity of their dendritic arbors. J Comp Neurol 390: 20–40

Riquelme R, Saldaña E, Osen KK, Ottersen OP, Merchán MA (2001) Colocalization of GABA and Glycine in the ventral nucleus of the lateral lemniscus in rat, an in situ hybridization and semiquantitative inmunocytochemical study. J Comp Neurol 432: 409–424

Rubio ME, Juiz JM (2004) Differential distribution of synaptic endings containing glutamate, glycine, and GABA in the rat dorsal cochlear nucleus. J Comp Neurol 477:253–272

Rubio ME, Gudsnuk KA, Smith Y, Ryugo DK (2008) Revealing the molecular layer of the primate dorsal cochlear nucleus. Neuroscience 154:99–113

Ryugo DK, Parks TN (2003) Primary innervation of the avian and mammalian cochlear nucleus. Brain Res Bull 60:435–456

Safieddine S, Eybalin M (1992) Triple immunofluorescence evidence for the coexistence of acetylcholine, enkephalins, and calcitonin gene-related peptide within efferent olivocochlear neurons in rats and Guinea-pigs. Eur J Neurosci 4:981–992

Safieddine S, Prior AM, Eybalin M (1997) Choline acetyltransferase, glutamate decarboxylase, tyrosine hydroxylase, calcitonin gene-related peptide and opioid peptides coexist in lateral efferent neurons of rat and Guinea-pig. Eur J Neurosci 9:356–367

Saldaña E, Merchán MA (1992) Intrinsic and commissural connections of the rat inferior colliculus. J Comp Neurol 319:417–437

Saldaña E, Feliciano M, 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

Schofield BR (1995) Projections from the cochlear nucleus to the superior paraolivary nucleus in Guinea pigs. J Comp Neurol 360:135–149

Schofield BR, Cant NB (1991) Organization of the superior olivary complex in the Guinea pig. I. Cytoarchitecture, cytochrome oxidase histochemistry, and dendritic morphology. J Comp Neurol 314: 645–670

Schreiner CE, Langner G (1988) Coding of temporal patterns in the central auditory nervous system. In: Edelman GM, Gall WE, Cowan WM (eds) Auditory function. Wiley, New York, pp 337–340

Schreiner CE, Langner G (1997) Laminar fine structure of frequency organization in auditory midbrain. Nature 388:383–386

Sherman SM, Guillery RW (1996) Functional organization of thalamocortical relays. J Neurophysiol 76: 1367–1395

Sinex DG, Lopez DE, Warr WB (2001) Electrophysiological responses of cochlear root neurons. Hear Res 370: 1–11

Sivaramakrishnan S, Oliver DL (2001) Distinct K^+ currents result in physiologically distinct cell types in the inferior colliculus of the rat. J Neurosci 21:2861–2877

Slepecky NB (1996) Structure of the mammalian cochlea. In: Dallos P, Popper AN, Fay RR (eds) The cochlea, Springer handbook of auditory research. Springer, New York, pp 44–129

Smith PH (1995) Structural and functional differences distinguish principal from nonprincipal cells in the Guinea pig MSO slice. J Neurophysiol 73:1653–1667

Smith PH, Rhode WS (1989) Structural and functional properties distinguish two types of multipolar cells in the ventral cochlear nucleus. J Comp Neurol 282: 595–616

Smith PH, Joris PX, Carney LH, Yin TCT (1991) Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. J Comp Neurol 304:387–407

Sommer I, Lingenhöhl K, Friauf E (1993) Principal cells of the rat medial nucleus of the trapezoid body, an intracellular in vivo study of their physiology and morphology. Exp Brain Res 95:223–239

Taranda J, Maison SF, Ballestero JA, Katz E, Savino J, Vetter DE, Boulter J, Liberman MC, Fuchs PA, Elgoyhen AB (2009) A point mutation in the hair cell nicotinic cholinergic receptor prolongs cochlear inhibition and enhances noise protection. PLoS Biol. https://doi.org/10.1371/journal.pbio.1000018

Thompson AM, Schofield BR (2000) Afferent projections of the superior olivary complex. Microsc Res Tech 51: 330–354

Vetter DE, Mugnaini E (1992) Distribution and dendritic features of three groups of rat olivocochlear neurons. A study with two retrograde cholera toxin tracers. Anat Embryol 185:1–16

Vetter DE, Saldaña E, 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

Wallace MN, Harper MS (1997) Callosal connections of the ferret primary auditory cortex. Exp Brain Res 116: 367–374

Warr WB (1992) Organization of olivocochlear efferent systems in mammals. In: Webster DB, Popper AN, Fay RR (eds) The mammalian auditory pathway, neuroanatomy. Springer, Berlin, pp 410–448

Warr WB, Boche JB, Neely ST (1997) Efferent innervation of the inner hair cell region, origins and terminations of two lateral olivocochlear systems. Hear Res 108: 89–111

Weedman DL, Ryugo DK (1996a) Projections from auditory cortex to the cochlear nucleus in rats, synapses on granule cell dendrites. J Comp Neurol 371:311–324

Weedman DL, Ryugo DK (1996b) Pyramidal cells in primary auditory cortex project to cochlear nucleus in rat. Brain Res 706:97–102

Weedman DL, Pongstaporn T, Ryugo DK (1996) Ultrastructural study of the granule cell domain of the cochlear nucleus in rats: mossy fiber endings and their targets. J Comp Neurol 369:345–360

White JS, Warr WB (1983) The dual origins of the olivocochlear bundle in the albino rat. J Comp Neurol 219:203–214

Wickesberg RE, Oertel D (1988) Tonotopic projection from the dorsal to the anteroventral cochlear nucleus of mice. J Comp Neurol 268:389–399

Winer JA (1985) The medial geniculate body of the cat. Adv Anat Embryol Cell Biol 86:1–97

Winer JA (1992) The functional architecture of the medial geniculate body and the primary auditory cortex. In: Webster DB, Popper AN, Fay RR (eds) The mammalian auditory pathway, neuroanatomy. Springer, New York, pp 222–409

Winer JA (2006) Decoding the auditory corticofugal systems. Hear Res 212:1–8

Winer JA, Larue DT (1988) Anatomy of glutamic acid decarboxylase immunoreactive neurons and axons in the rat medial geniculate body. J Comp Neurol 278: 47–68

Winer JA, Lee CC (2007) The distributed auditory cortex. Hear Res 229:3–13

Winer JA, Morest DK (1983) The medial division of the medial geniculate body of the cat: implications for thalamic organization. J Neurosci 3:2629–2651

Winer JA, Schreiner CE (2011) The auditory cortex. Springer, New York

Winer JA, Kelly JB, Larue DT (1999) Neural architecture of the rat medial geniculate body. Hear Res 31:19–41

Wu SH (1999) Physiological properties of neurons in the ventral nucleus of the lateral lemniscus of the rat, intrinsic membrane properties and synaptic responses. J Neurophysiol 81:2862–2874

Ye Y, Machado DG, Kim DO (2000) Projection of the marginal shell of the anteroventral cochlear nucleus to olivocochlear neurons in the cat. J Comp Neurol 420: 127–138

Young ED, Davis KA (2002) Circuitry and function of the dorsal cochlear nucleus, chapter 5. In: Oertel D, Fay RR, Popper AN (eds) Springer handbook of auditory research, Integrative functions in the mammalian auditory pathway, vol 15. Springer, New York, pp 160–206

Young ED, Shofner WP, White JA, Robert JM, Voigt HF (1988) Response properties of cochlear nucleus neurons in relationship to physiological mechanisms. In: Edelman GM, Gall WE, Cowan WM (eds) Auditory function, neurobiological bases of hearing. Wiley, New York, pp 277–312

Yu XJ, Xu XX, He S, He J (2009) Change detection by thalamic reticular neurons. Nat Neurosci 12:1165–1170

Zhang H, Kelly JB (2006a) Responses of neurons in the rat's ventral nucleus of the lateral lemniscus to monaural and binaural tone bursts. J Neurophysiol 95: 2501–2512
Zhang H, Kelly JB (2006b) Responses of neurons in the rat's ventral nucleus of the lateral lemniscus to amplitude-modulated tones. J Neurophysiol 96: 2905–2914
Zhang Y, Wu SH (2000) Long-term potentiation in the inferior colliculus studied in rat brain slice. Hear Res 1(47):92–103
Zhao M, Wu SH (2001) Morphology and physiology of neurons in the ventral nucleus of the lateral lemniscus in rat brain slices. J Comp Neurol 433: 255–271
Zhou J, Shore S (2006) Convergence of spinal trigeminal and cochlear nucleus projections in the inferior colliculus of the Guinea pig. J Comp Neurol 495: 100–112

Anesthesia, Neural Population Models of

D. Alistair Steyn-Ross[1], Moira Steyn-Ross[1] and Jamie Sleigh[2]
[1]School of Engineering, University of Waikato, Hamilton, New Zealand
[2]Waikato Clinical School, University of Auckland, Waikato Hospital, Hamilton, New Zealand

Definition

General anesthesia is a reversible, drug-induced state of unconsciousness characterized by lack of awareness of surroundings, lack of responsiveness to painful stimuli (nociception), and inability to form memories (amnesia). The change in brain state from wakeful to unconscious produces alterations in cortical electrical activity that can be monitored with electrodes placed on the scalp (electroencephalogram (EEG)) or on the surface of the cortex (electrocorticogram (ECoG)). The goal of neural modelers is to develop equations that describe the gross behavior of spatially averaged populations of neurons during both induction of and recovery from general anesthesia.

Detailed Description

Classes of General Anesthesia

There are two broad classes of anesthetic drugs: inductive agents (such as propofol, etomidate, isoflurane) that produce a slowed sleeplike EEG and dissociative agents (e.g., ketamine, nitrous oxide) that induce a dissociated state with an activated EEG similar to that of REM sleep.

Most commonly used intravenous and volatile agents – such as propofol or sevoflurane – boost inhibition by increasing the influx of chloride ions at gamma-aminobutyric acid (GABA) receptors on postsynaptic membranes (Weir 2006), causing the postsynaptic neuron to become hyperpolarized. In contrast, dissociative drugs are believed to disrupt excitatory synaptic transmission. In both cases, the excitatory – inhibitory balance required for normal brain function has been shifted to favor inhibition.

The Induction: Recovery Trajectory

At low concentrations, most GABAergic agents (e.g., propofol, sevoflurane, etomidate) cause a paradoxical boost in cortical activity (called the "biphasic effect") across most EEG frequency bands (Kuizenga et al. 2001), with the biphasic peak appearing first in the high beta frequencies (24–28 Hz), then sliding smoothly towards lower frequencies in time (e.g., see Fig. 3 of Koskinen et al. (2005)). With further increase in concentration, the EEG slows as large-amplitude delta-band oscillations (1–4 Hz) become dominant, then changes to an intermittent burst–suppression pattern (bursting activity alternating with relative silence), and finally collapses into a flat-line trace at the deepest levels of comatose anesthesia.

This sequence is reversed as the anesthetic drug is eliminated naturally from the body, allowing the patient to return to consciousness. However, the fact that the recovery of responsiveness generally occurs at a *lower* drug concentration (as measured in the blood) than that required to induce unresponsiveness suggests a hysteresis separation between induction and recovery trajectories. Part of this hysteresis can be explained in terms of the time required for the drug to diffuse

across the blood–brain barrier (Voss et al. 2007) and so can be compensated using pharmacokinetics models (Roberts 2007), but such compensations are typically only partially successful (Ludbrook et al. 1999; Coppens et al. 2010). The remaining hysteresis may be a consequence of a recently proposed "neural inertia" that resists transitions between conscious and unconscious states (Friedman et al. 2010); such distinct induction/recovery paths arise naturally if the brain has access to multiple steady states as suggested by the modeling of Steyn-Ross et al. (1999, 2004).

Cellular Effects of General Anesthetic Drugs

Studies of propofol, halothane, and isoflurane have shown that, at drug concentrations rendering human subjects unresponsive, cerebral blood flow and metabolism are reduced by about 50% (Antkowiak 2002) as a result of global reductions in cortical activity. This is consistent both with in vivo investigations in rat cortex – where sedative-level concentrations were found to suppress neural firing rates by 50–70% (Gaese and Ostwald 2001) – and with cultured brain-slice studies in which low concentrations of general anesthetics (GABAergic agonists propofol, halothane, isoflurane, enflurane, sevoflurane, etomidate, ethanol, and pentobarbital and the non-GABAergic agent ketamine) significantly decreased mean firing rates (Antkowiak 2002).

All anesthetic drugs influence cellular function in a number of different ways, but the major mechanism for GABAergic suppression of firing rates is believed to be the prolongation of the opening of chloride channels on the postsynaptic neuron, thus causing a substantial increase in negative charge transfer (by a factor of 2–4 times control at clinically relevant concentrations (Kitamura et al. 2003; Banks and Pearce 1999)) during the inhibitory postsynaptic current (IPSC) pulse.

Dissociative drugs reduce excitatory transmission by blocking *N*-methyl-d-aspartate (NMDA) glutamate channels, which probably has a significant role in producing the characteristic dissociated anesthetic state (Petrenko et al. 2013); however, these drugs also have other effects such as inhibition of hyperpolarization-activated cyclic nucleotide-gated (HCN1) channels (Chen et al. 2009) or increased potassium channel opening (Gruss et al. 2004).

Modeling Anesthetic Effects

The challenge for anesthesia modelers is to bridge the scales from the microscopic cellular drug effects to the consequent *macroscopic* population behaviors detected with scalp or cortical electrodes. By considering spatially averaged ("mean-field") properties of cortical tissue, we can avoid the need (and computational expense) of attempting to explicitly represent myriads of individual neurons (as is done in neural networks). There is a steadily growing interest in applying mean-field methods to the challenge of understanding anesthesia; see Foster et al. (2008) and Steyn-Ross et al. (2011) for reviews.

The notion of *neural fields* dates from foundation work by Wilson and Cowan (1972) that modeled the brain as homogenous populations of excitatory and inhibitory neurons. The first attempt at modeling propofol anesthesia by Steyn-Ross et al. (1999) incorporated prolongation of inhibitory response into the mean-field neural model of Liley et al. (1999); it predicted the possibility of multiple steady states with distinct first-order phase transitions between activated ("conscious") and inactivated ("unconscious") states and provided a possible explanation for the hysteretically separated biphasic power surges observed at loss and recovery of consciousness (Kuizenga et al. 2001).

Subsequent work by Bojak and Liley (2005) on isoflurane anesthesia showed that, for suitable choices of cortical parameters, a smooth descent into unconsciousness can also generate a biphasic drug response. Using an alternative mean-field model, Hutt and colleagues (Hutt and Schimansky-Geier 2008; Hutt and Longtin 2010) predicted that biphasic power surges can be expected for both the bistable (jump transition) and monostable (smooth) inductions of anesthesia.

General anesthetic agents are widely used to treat seizures, but paradoxically, some anesthetics (e.g., enflurane) can also provoke cortical seizures when the patient is deeply anesthetized. Liley and

Bojak (2005) and Wilson et al. (2006) used mean-field modeling to show that subtle changes in the shape and duration of the drug-induced inhibitory postsynaptic response can explain why enflurane, but not isoflurane, is seizurogenic.

An important part of general anesthesia is the suppression of noxious stimuli. A practical index of antinociception has been developed from a mean-field model (Liley et al. 2010) that informed construction of an autoregressive – moving-average (ARMA) noise-driven filter whose output approximates the scalp-recorded EEG. The mean filter frequency tracks the level of propofol-induced hypnosis ("cortical state"), while the decrease in required noise intensity ("cortical input") tracks the concentration of a coadministered analgesic agent (remifentanil). This computed "cortical input" signal is presumed to be a measure of cortical stimulus, both noxious and normal, entering from the thalamus, and potentially allows differentiation between hypnotic and analgesic drug effects.

The unconscious state of anesthesia and of deepest natural sleep are both characterized by large-amplitude, slow (0.5–4 Hz) delta waves of EEG activity. The source of these slow waves is unknown but is generally supposed to originate from gradual alternations in depolarizing and hyperpolarizing ionic currents. By introducing a slow ionic gating variable into a mean-field model for desflurane anesthesia, Molaee-Ardekani et al. (2007) demonstrated emergence of realistic slow waves. A quite different slow-wave mechanism has been proposed by Steyn-Ross et al. (2013): if inhibitory gap junctions are included in the two-dimensional cortical sheet, then a Turing (pattern-forming) instability can interact with a weakly damped low-frequency Hopf instability to produce turbulent slow-wave activity across the cortex. Anesthetic-induced closure of inhibitory gap junctions (Wentlandt et al. 2006) is predicted to weaken the Turing instability in favor of the Hopf oscillation.

There has been interest in modeling some of the specific details of EEG spectral changes caused by various general anesthetic drugs, for example, the displacement in alpha peak frequency induced by ketamine (Bojak et al. 2013) or propofol (Hindriks and van Putten 2012; Hutt 2013) and the burst–suppression pattern of deep anesthesia (Liley and Walsh 2013).

Increasingly there has been a realization that general anesthesia may disrupt neuronal networks in an anatomically specific fashion (Kuhlmann et al. 2013; Lee et al. 2013) and that the current homogenous and isotropic neuronal population models might need to include aspects of network topology. This has led to attempts to link EEG patterns probabilistically with underlying anesthetic effects on inhibitory and excitatory neuronal groups – this should provide a quantitative basis for the estimation of model parameters. At an abstract level, dynamic causal-modeling methods have been employed (Moran et al. 2011; Boly et al. 2012), but a more direct Bayesian approach – which has been used for natural sleep (Dadok et al. 2013) – could be applied to anesthesia EEG.

Cross-References

- ▶ Bifurcations, Neural Population Models and
- ▶ Epilepsy, Neural Population Models of
- ▶ Gap Junctions in Small Networks
- ▶ Gap Junctions, Neural Population Models and
- ▶ Neural Population Model
- ▶ Pattern Formation in Neural Population Models
- ▶ Phase Transitions, Neural Population Models and
- ▶ Sleep, Neural Population Models of

References

Antkowiak B (2002) In vitro networks: cortical mechanisms of anaesthetic action. Br J Anaesth 89(1): 102–111

Banks MI, Pearce RA (1999) Dual actions of volatile anesthetics on GABA(a) IPSCs: dissociation of blocking and prolonging effects. Anesthesiology 90(1):120–134

Bojak I, Liley DT (2005) Modeling the effects of anesthesia on the electroencephalogram. Phys Rev E 71(4 Pt 1):041902. http://www.ncbi.nlm.nih.gov/pubmed/15903696

Bojak I, Day HC, Liley DT (2013) Ketamine, propofol, and the EEG: a neural field analysis of HCN1-mediated

interactions. Front Comput Neurosci 7:22. https://doi.org/10.3389/fncom.2013.00022. http://www.ncbi.nlm.nih.gov/pubmed/23576979

Boly M, Moran R, Murphy M, Boveroux P, Bruno MA, Noirhomme Q, Ledoux D, Bonhomme V, Brichant JF, Tononi G, Laureys S, Friston K (2012) Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. J Neurosci 32(20):7082–7090. https://doi.org/10.1523/JNEUROSCI.3769-11.2012

Chen X, Shu S, Bayliss DA (2009) HCN1 channel subunits are a molecular substrate for hypnotic actions of ketamine. J Neurosci 29(3):600–609. https://doi.org/10.1523/JNEUROSCI.3481-08.2009

Coppens M, Van Limmen JGM, Schnider T, Wyler B, Bonte S, Dewaele F, Struys MMRF, Vereecke HEM (2010) Study of the time course of the clinical effect of propofol compared with the time course of the predicted effect-site concentration: performance of three pharmacokinetic-dynamic models. Br J Anaesth 104(4):452–458. https://doi.org/10.1093/bja/aeq028

Dadok VM, Kirsch HE, Sleigh JW, Lopour BA, Szeri AJ (2013) A probabilistic framework for a physiological representation of dynamically evolving sleep state. J Comput Neurosci. https://doi.org/10.1007/s10827-013-0489-x

Foster BL, Bojak I, Liley DTJ (2008) Population based models of cortical drug response: insights from anaesthesia. Cogn Neurodyn 2(4):283–296. https://doi.org/10.1007/s11571-008-9063-z

Friedman EB, Sun Y, Moore JT, Hung HT, Meng QC, Perera P, Joiner WJ, Thomas SA, Eckenho RG, Sehgal A, Kelz MB (2010) A conserved behavioral state barrier impedes transitions between anesthetic-induced unconsciousness and wakefulness: evidence for neural inertia. PLoS One 5(7):e11903. https://doi.org/10.1371/journal.pone.0011903

Gaese BH, Ostwald J (2001) Anesthesia changes frequency tuning of neurons in the rat primary auditory cortex. J Neurophysiol 86(2):1062–1066

Gruss M, Bushell TJ, Bright DP, Lieb WR, Mathie A, Franks NP (2004) Two-pore-domain K+ channels are a novel target for the anesthetic gases xenon, nitrous oxide, and cyclopropane. Mol Pharmacol 65(2):443–452. https://doi.org/10.1124/mol.65.2.443

Hindriks R, van Putten MJAM (2012) Meanfield modeling of propofol-induced changes in spontaneous EEG rhythms. NeuroImage 60(4):2323–2334. https://doi.org/10.1016/j.neuroimage.2012.02.042

Hutt A (2013) The anesthetic propofol shifts the frequency of maximum spectral power in EEG during general anesthesia: analytical insights from a linear model. Front Comput Neurosci 7:2. https://doi.org/10.3389/fncom.2013.00002. http://www.ncbi.nlm.nih.gov/pubmed/23386826

Hutt A, Longtin A (2010) Effects of the anesthetic agent propofol on neural populations. Cogn Neurodyn 4(1):37–59. https://doi.org/10.1007/s11571-009-9092-2

Hutt A, Schimansky-Geier L (2008) Anesthetic-induced transitions by propofol modeled by nonlocal neural populations involving two neuron types. J Biol Phys 34(3–4):433–440. https://doi.org/10.1007/s10867-008-9065-4

Kitamura A, Marszalec W, Yeh JZ, Narahashi T (2003) Effects of halothane and propofol on excitatory and inhibitory synaptic transmission in rat cortical neurons. J Pharmacol Exp Ther 304(1):162–171. https://doi.org/10.1124/jpet.102.043273

Koskinen M, Mustola S, Seppänen T (2005) Relation of EEG spectrum progression to loss of responsiveness during induction of anesthesia with propofol. Clin Neurophysiol 116(9):2069–2076. https://doi.org/10.1016/j.clinph.2005.06.004

Kuhlmann L, Foster BL, Liley DT (2013) Modulation of functional EEG networks by the NMDA antagonist nitrous oxide. PLoS One 8(2):e56434. https://doi.org/10.1371/journal.pone.0056434. http://www.ncbi.nlm.nih.gov/pubmed/23457568

Kuizenga K, Wierda JM, Kalkman CJ (2001) Biphasic EEG changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane. Br J Anaesth 86(3):354–360

Lee U, Ku S, Noh G, Baek S, Choi B, Mashour GA (2013) Disruption of frontal-parietal communication by ketamine, propofol, and sevoflurane. Anesthesiology 118(6):1264–1275. https://doi.org/10.1097/ALN.0b013e31829103f5

Liley DTJ, Bojak I (2005) Understanding the transition to seizure by modeling the epileptiform activity of general anesthetic agents. Clin Neurophysiol 22(5):300–313

Liley DT, Walsh M (2013) The mesoscopic modeling of burst suppression during anesthesia. Front Comput Neurosci 7:46. https://doi.org/10.3389/fncom.2013.00046. http://www.ncbi.nlm.nih.gov/pubmed/23641211

Liley DTJ, Cadusch PJ, Wright JJ (1999) A continuum theory of electro-cortical activity. Neurocomputing 26–27:795–800

Liley DT, Sinclair NC, Lipping T, Heyse B, Vereecke HE, Struys MM (2010) Propofol and remifentanil differentially modulate frontal electroencephalographic activity. Anesthesiology 113(2):292–304. https://doi.org/10.1097/ALN.0b013e3181e3d8a6

Ludbrook GL, Upton RN, Grant C, Martinez A (1999) Prolonged dysequilibrium between blood and brain concentrations of propofol during infusions in sheep. Acta Anaesthesiol Scand 43(2):206–211

Molaee-Ardekani B, Senhadji L, Shamsollahi MB, Vosoughi-Vahdat B, Wodey E (2007) Brain activity modeling in general anesthesia: enhancing local mean-field models using a slow adaptive firing rate. Phys Rev E 76(4 Pt 1):041911. http://www.ncbi.nlm.nih.gov/pubmed/17995030

Moran RJ, Jung F, Kumagai T, Endepols H, Graf R, Dolan RJ, Friston KJ, Stephan KE, Tittge-meyer M (2011) Dynamic causal models and physiological inference: a

validation study using isoflurane anaesthesia in rodents. PLoS One 6(8):e22790. https://doi.org/10.1371/journal.pone.0022790. http://www.ncbi.nlm.nih.gov/pubmed/21829652
Petrenko AB, Yamakura T, Sakimura K, Baba H (2013) Defining the role of NMDA receptors in anesthesia: are we there yet? Eur J Pharmacol 723C:29–37. https://doi.org/10.1016/j.ejphar.2013.11.039
Roberts F (2007) Pharmacokinetics and anaesthesia. Contin Educ Anaesth Crit Care Pain 7(1):25–29. https://doi.org/10.1093/bjaceaccp/mkl058
Steyn-Ross ML, Steyn-Ross DA, Sleigh JW, Liley DTJ (1999) Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition. Phys Rev E 60(6 Pt B):7299–7311. http://www.ncbi.nlm.nih.gov/pubmed/11970675
Steyn-Ross ML, Steyn-Ross DA, Sleigh JW (2004) Modelling general anaesthesia as a first-order phase transition in the cortex. Prog Biophys Mol Biol 85(2–3):369–385. https://doi.org/10.1016/j.pbiomolbio.2004.02.001
Steyn-Ross DA, Steyn-Ross ML, Sleigh JW, Wilson MT (2011) Progress in modeling EEG effects of general anesthesia: biphasic response and hysteresis. In: Hutt A (ed) Sleep and anesthesia: neural correlates in theory and experiment, chapter 8, Springer series in computational neuroscience, vol 15. Springer, New York, pp 167–194. https://doi.org/10.1007/978-1-4614-0173-5n8
Steyn-Ross ML, Steyn-Ross DA, Sleigh JW (2013) Interacting Turing-Hopf instabilities drive symmetry-breaking transitions in a mean-field model of the cortex: a mechanism for the slow oscillation. Phys Rev X 3(2):021005. https://doi.org/10.1103/PhysRevX.3.021005. http://link.aps.org/doi/10.1103/PhysRevX.3.021005
Voss LJ, Ludbrook G, Grant C, Upton R, Sleigh JW (2007) A comparison of pharmacokinetic/pharmacodynamic versus mass-balance measurement of brain concentrations of intra-venous anesthetics in sheep. Anesth Analg 104(6):1440–1446. https://doi.org/10.1213/01.ane.0000263274.62303.1a
Weir CJ (2006) The molecular mechanisms of general anaesthesia: dissecting the GABAA receptor. Contin Educ Anaesth Crit Care Pain 6(2):49–53. https://doi.org/10.1093/bjaceaccp/mki068
Wentlandt K, Samoilova M, Carlen PL, El Beheiry H (2006) General anesthetics inhibit gap junction communication in cultured organotypic hippocampal slices. Anesth Analg 102(6):1692–1698. https://doi.org/10.1213/01.ane.0000202472.41103.78
Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J 12:1–24
Wilson MT, Sleigh JW, Steyn-Ross DA, Steyn-Ross ML (2006) General anesthetic-induced seizures can be explained by a mean-field model of cortical dynamics. Anesthesiology 104:588–593

Animal Calls

▶ Pulse-Resonance Sounds

Animal Models of Pain

▶ Biomechanical Model of Low Back Pain

Anomalous Rectifier Potassium Channels

▶ Inward Rectifier Potassium Channels

Anti-Hebbian Learning

Yoonsuck Choe
Department of Computer Science and Engineering, Texas A&M University, College Station, TX, USA

Definition

Anti-Hebbian learning is a form of activity-dependent synaptic plasticity that is defined as the opposite of Hebbian learning. Hebbian learning is commonly defined as follows: correlated activation in the pre- and postsynaptic neurons leading to the strengthening of the connection between the two neurons. However, the original definition offered in Hebb (1949) talks about the increase in the presynaptic neuron's efficiency of eliciting activity in the postsynaptic neuron, under the same correlated firing condition. These two definitions (strengthening of connection vs. increased efficiency) are compatible only when the presynaptic neuron is excitatory. They are contradictory when the presynaptic neuron is inhibitory: increased connection strength (weight, efficacy) corresponds to decreased efficiency in

eliciting response. Assuming the original definition of Hebbian learning, we can define anti-Hebbian learning as a form of synaptic plasticity where correlated activation in the pre- and postsynaptic neurons leads to the reduction in the efficiency of the presynaptic neuron's ability to elicit activation of the postsynaptic neuron.

Detailed Description

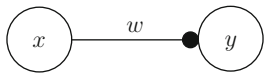
In this entry, we will review computational formulations of anti-Hebbian learning and their theoretical implications. We will also briefly touch upon neurobiological correlates of the learning mechanism. The literature on anti-Hebbian learning is not as rich as that on Hebbian learning. A review of Hebbian learning by Frégnac (2003) includes a brief discussion on the neurobiological bases of anti-Hebbian learning. See Földiák (1990) for a computational treatment of the subject and Palmieri et al. (1993) for a review.

Basic Formulation

Consider a simple configuration with a single presynaptic neuron x connecting to a single postsynaptic neuron y with synaptic weight w as shown above (Fig. 1). Anti-Hebbian learning is the same as Hebbian learning, except for the flip of the sign.

$$\frac{dw}{dt} = -\eta xy, \tag{1}$$

where the learning rate η is a small, fixed, positive value. The neuronal activities x and y are normally assumed to be positive (or zero), but for purely computational purposes, they can be negative values as well. In Hebbian learning, adaptive threshold or normalization approaches are used to check unbounded growth in the synaptic weight. However, in anti-Hebbian learning, such a check is unnecessary since the mechanism is inherently stable (Földiák 1990). At first glance, it appears that the weight w can tend toward $-\infty$, but this will never happen because high levels of inhibition will shut off y after a certain point and thus w will not change any further thereafter. When x and y are allowed to be negative, w can increase under the same learning rule. Some formulations prevent this from causing infinite growth by imposing a limit (w is assumed to be negative and it is reset to 0 as soon as it becomes positive; Földiák 1990), but others do not include such a limit (e.g., Girolami and Fyfe 1996).

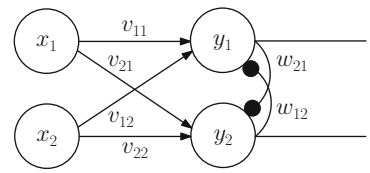


Anti-Hebbian Learning, Fig. 1 A pair of neurons forming an anti-Hebbian connection

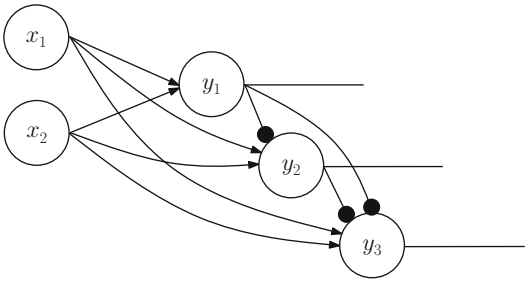
Theoretical Perspectives

Anti-Hebbian learning is usually combined with Hebbian learning to produce interesting theoretical and practical results. Figure 2 below shows such an example (adapted from Földiák 1990). In this figure, two downstream neurons y_1 and y_2 receive afferent input from x_1 and x_2 through Hebbian synapses (with weights v_{ij}) and exchange activations via anti-Hebbian lateral connections (with weights w_{ij}). (The subscripts ij on the weights indicate the target [i] and source neuron index [j], respectively.)

Földiák showed that the above network can learn to decorrelate the output neurons' activity (sparse coding in neurons y_i), which results in the learning of afferent representations (v_{ij}) that are components that make up the input mixtures. The model included an anti-Hebbian rule similar to Eq. 1 with an added threshold (w_{ij}), a Hebbian



Anti-Hebbian Learning, Fig. 2 Neurons connected with Hebbian (*arrows*) and anti-Hebbian synapses (*discs*). (Simplified from Földiák 1990)



Anti-Hebbian Learning, Fig. 3 A hierarchy of neurons y_i with Hebbian afferent connections (*arrows*) and anti-Hebbian connections in a cascade (*discs*). (Adapted from Carlson 1990)

rule with weight decay (v_{ij}), and a dynamic neural activation equation $\left(\frac{dy_i}{dt}\right)$ with adaptive threshold and sigmoid nonlinearity (for details, see Földiák 1990).

Now consider a hierarchical network shown above (Fig. 3). Hebbian learning is known to extract the first principal component of the inputs (Oja 1982), so neuron y_1 would serve this role. Using a hierarchical network like Fig. 3 with anti-Hebbian connections arranged in a cascade, second, third, and subsequent principal components of the input can be found (see Carlson 1990 for a review and details). The main idea is that once y_i learned the i-th principal component, and y_{i+1} is decorrelated with y_1, y_2, ..., y_i through anti-Hebbian connections (note that y_3 receives anti-Hebbian connections from both y_1 and y_2), y_{i+1} will find the $i + 1$-th principal component.

For a review of more complicated network topologies that involve anti-Hebbian learning, see Palmieri et al. (1993).

Neurobiological Underpinnings

As Hebbian learning is usually associated with long-term potentiation (LTP), anti-Hebbian learning is typically explained by long-term depression (LTD). LTP is a phenomenon where high frequency stimulation of the presynaptic neuron leads to a prolonged increase in synaptic efficacy (Bliss and Collingridge 1993). LTD is the opposite, where low-frequency stimulation causes a prolonged decrease in synaptic efficacy (Dudek and Bear 1992); thus, it fits the anti-Hebbian profile.

Examples of LTD associated with anti-Hebbian mechanisms have been found in the mammalian cerebellum (e.g., mouse and rat) and also in the electrosensory lobe (ELL) in the teleost electric fishes, and these LTD mechanisms are thought to be playing an anti-Hebbian role (see Frégnac 2003 for a brief review).

In some cases, LTP can also be associated with a form of anti-Hebbian learning, when LTP is induced not by correlated activation but by presynaptic firing not being met by or negated by the postsynaptic activity. Such a phenomenon has been observed in the hippocampus and has been dubbed anti-Hebbian LTP. Kullmann and Lamsa (2007) provide an extensive review on this topic and provide a comparison of anti-Hebbian LTD and anti-Hebbian LTP.

Finally, spike timing-dependent plasticity (STDP; see Caporale and Dan 2008 for a review) has also been linked to anti-Hebbian learning. STDP is a short-term synaptic plasticity mechanism where depending on the ordering of pre- and postsynaptic events, either LTP or LTD can entail. LTP is induced when the presynaptic activity precedes postsynaptic activity and LTD when the ordering is reversed. Some views the LTD part of STDP as implementing an anti-Hebbian rule (see Nelson 2004 for a discussion).

Applications of Anti-Hebbian Learning

Anti-Hebbian learning has been applied to several signal and data processing tasks including vision and speech processing. Girolami and Fyfe (1996) used anti-Hebbian rule to learn finite impulse response (FIR) filter coefficients and applied the technique to blind source separation in the speech recognition domain. Schraudolph and Sejnowski (1992) used anti-Hebbian rule to learn invariances in disparity tuning for stereo vision. Földiák (1990) combined Hebbian and anti-Hebbian rule to learn sparse representations from overlapping visual inputs.

Cross-References

- ▶ Hebbian Learning
- ▶ Spike-Timing Dependent Plasticity, Learning Rules

References

Bliss TVP, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31–39

Caporale N, Dan Y (2008) Spike timing-dependent plasticity: a Hebbian learning rule. Ann Rev Neurosci 31: 25–46

Carlson A (1990) Anti-Hebbian learning in a non-linear neural network. Biol Cybern 64:171–176

Dudek SM, Bear MF (1992) Homosynaptic long-term depression in area CA1 of hippocampus and effects of *N*-methyl-d-aspartate receptor blockade. Proce Natl Acad Sci USA 89:4363–4367

Földiák P (1990) Forming sparse representations by local anti-Hebbian learning. Biol Cybern 64: 165–170

Frégnac Y (2003) Hebbian synaptic plasticity. In: Arbib MA (ed) The handbook of brain theory and neural networks, vol 2. MIT press, Cambridge, MA, pp 515–522

Girolami M, Fyfe C (1996) A temporal model of linear anti-Hebbian learning. Neural Process Lett 4: 139–148

Hebb DO (1949) The organization of behavior: a neuropsychological theory. Wiley, New York

Kullmann DM, Lamsa KP (2007) Long-term synaptic plasticity in hippocampal interneurons. Nat Rev Neurosci 8:687–699

Nelson SB (2004) Hebb and anti-Hebb meet in the brainstem. Nat Neurosci 7:687–688

Oja E (1982) A simplified neuron model as a principal component analyzer. J Math Biol 15:267–273

Palmieri F, Zhu J, Chang C (1993) Anti-Hebbian learning in topologically constrained linear networks: a tutorial. IEEE Trans Neural Netw 4:748–761

Schraudolph NN, Sejnowski TJ (1992) Competitive anti-Hebbian learning of invariants. In: Moody JE, Hanson SJ, Lippmann RP (eds) Advances in neural information processing systems. Morgan Kaufmann, San Mateo, pp 1017–1024

Aperture Problem

- ▶ Somatosensory Cortex: Neural Coding of Motion

Application of Declarative Programming in Neurobiology

Thomas J. Anastasio
Department of Molecular and Integrative Physiology, and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Definition

The main technique in computational neuroscience is imperative programming, which is often used to implement simulations of dynamics, and describes *how* a computation is performed. A complement to this approach is declarative programming. Declarations provide descriptions of relationships between elements, effectively describing *what* the computation should accomplish. The use of declarative programming for modeling biological processes is still in its infancy (Fisher and Henzinger 2007) yet has shown itself to be a valuable first step for analysis of the seemingly impenetrable complexity of molecular interactomics: the interplay of the myriad proteins and signaling species in the cell. Declarative programming can also be used at the connectomic level of understanding connections among neurons or among brain areas.

Detailed Description

Declarative Programming

The *declarations* of a declarative program describe the relationships between system elements. Because of this descriptive nature, a model implemented in a declarative program may be considered as a system *specification*. Many declarative programming languages utilize term-rewriting logic: the declarations are rewrite laws that specify how one term should be replaced by another. Abstract rewriting systems are written with arrows, specifying that the left-hand side (LHS) is to be replaced by the term

on the right-hand side (RHS). Other notations, such as Backus-Naur Form, utilize some variant of an equal sign to specify that RHS is replaced by LHS.

Crucially, because the declarations are descriptions of system properties, the specification can be used not only for simulation but also for analyses such as state-space search and temporal-logic model checking (explained below; for a general reference see Huth and Ryan 2004). Term-rewriting declarative languages are readily used for expressing data-driven models constructed directly from experimental observations. These models do not require an overarching conceptual framework or preconceived hypothesis, allowing the enormous amounts of data generated by neurobiological experiments to be directly entered. This allows the data to "speak for itself" and provides a model that can be used as a tool to help develop subsequent hypotheses.

The first term-rewriting language was the λ-calculus, developed in the 1930s by Alonzo Church (Cardone and Hindley 2006). Since then several other term-rewriting, declarative programming languages have been developed. Some well-known examples are Scheme (http://www.r6rs.org/), Alloy (http://alloy.mit.edu/alloy/index.html), Simile (http://www.simulistics.com/tour/declarative.htm), and ECLiPSe (http://eclipseclp.org/index.html). One declarative programming language that has been used effectively for modeling molecular interactions and neurobiological processes is Maude (http://maude.cs.uiuc.edu/; Clavel et al. 2007).

A specification in Maude is based on an underlying algebra, which is defined by *sorts* (data types) and by *operations* that are allowed for each sort (e.g., natural numbers – sort, can be added together – operation). Preset algebras such as the natural numbers are available, but new algebras can be defined by the user. Declarations in Maude are written in terms of the underlying algebra and can be expressed either as equations or as rules. Equations in Maude specify functional relationships that simplify the state of the system being modeled, while rules specify transitional relationships that change the state of the model system. For example, an equation could specify that 2 apples plus 4 apples plus 3 apples equals 9 apples, while a rule could specify that 9 apples can transition to an apple pie. Equations and rules can be unconditional (`eq` and `rl`, respectively) or conditional (`ceq` and `crl`). For example, a conditional rule could specify that 9 apples can transition to an apple pie, but only if mom is available to bake it. While an unconditional equation or rule applies whenever its LHS is present, a conditional rule requires both the LHS and the required conditions.

All rewrite systems have rules (a.k.a. syntax) that transition the system state and evaluations (a.k.a. semantics) that can reduce the state, in whole or in part, to its underlying value at arbitrary times (Kain 1972). Another characteristic of Maude is that equations (i.e., evaluations) *must* execute whenever they apply while applicable rules may execute or not. Rules, by executing, can transition the system to a state in which equations become applicable that were not so in the previous state. In using Maude for simulation, all applicable rules would ultimately execute according to a *fairness criterion*: every applicable rule must execute once before any other rule can execute twice. In using Maude for analyses, such as state-space search or logical model checking, the rules execute in all possible combinations and orders. Because rules cause state transitions, Maude thereby elaborates the state transition tree implied by the specification and later searches this tree for specific states or to verify certain temporal-logic propositions.

Maude has been used to specify and analyze a broad range of complex engineered systems used for computer network security, encryption, and avionics, among others (Meseguer 2012). Maude is also used to model other programming languages, and analysis of a Maude model can verify that a language behaves as intended. This type of analytical capability suggests how Maude can be used to assess interactomic "programs," so as to determine where they have vulnerabilities, predicting sites for potential failure which would produce disease, as well as sites where intervention could compensate for these failures,

predicting sites where pharmacological intervention would be of value.

Declarative Models of Neurobiological and Neurodegenerative Processes

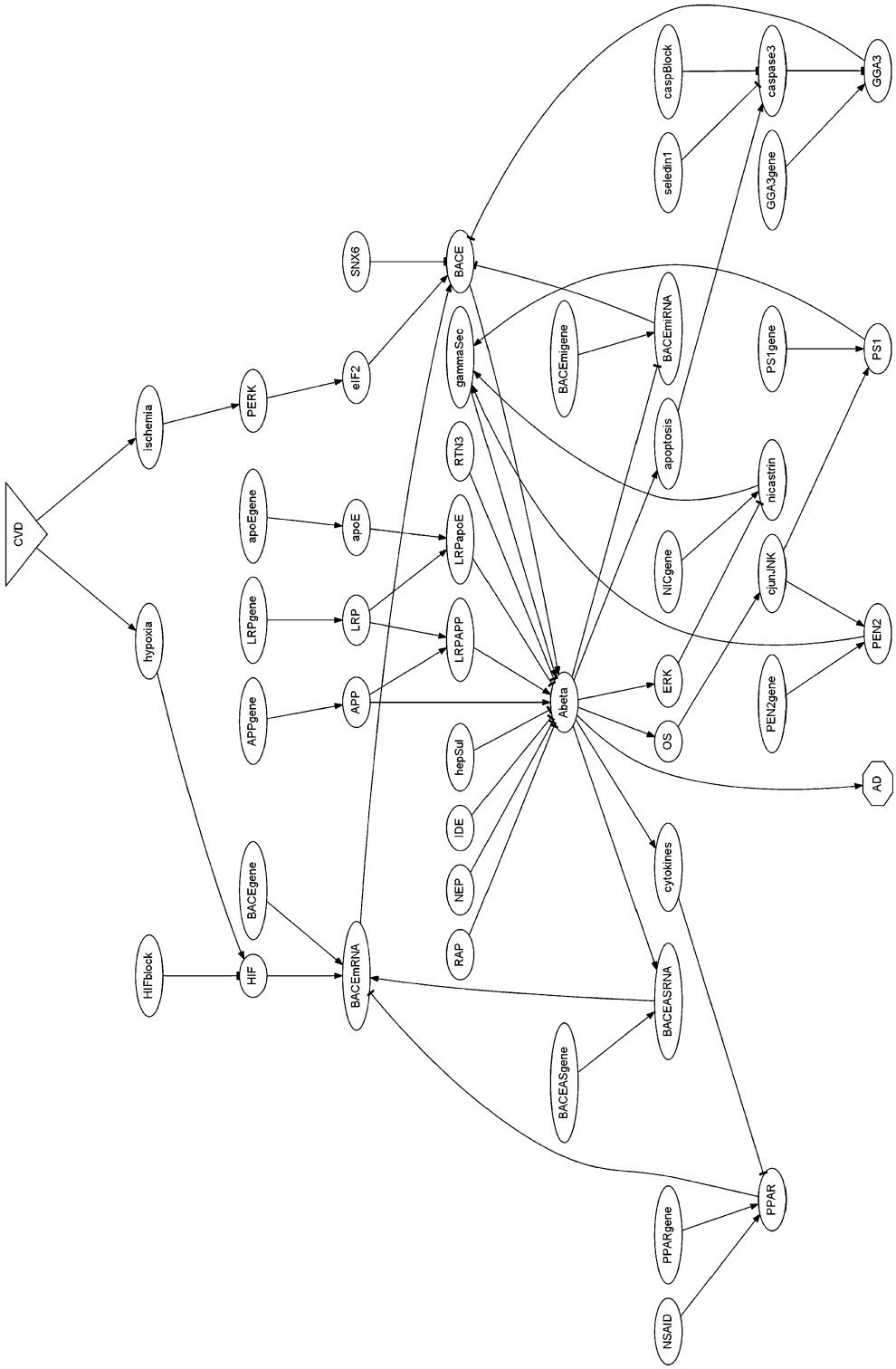
Maude has been used to model several biological processes (Eker et al. 2002; Talcott 2008). An initial neurobiological application explores the synaptic plasticity thought to underlie fear conditioning and its extinction (Anastasio 2013b). In that microconnectomic model, rules specified synaptic weight changes while equations specified changes in neural responses caused by weight changes. Thus, each state of the model system had a different one of the many possible configurations of synaptic weights. This particular model had only nine synapses, which were allowed only a limited number of discrete weights, so the state space was large but still small enough to permit exhaustive search. Analysis via state-space search provided testable hypotheses by revealing which of these very many synaptic weight configurations were compatible with fear conditioning followed by extinction.

More recent neurobiological applications involve the food-intake control system (Tabe-Bordbar and Anastasio 2016) and the monoaminergic neurotransmitter system (Camacho and Anastasio 2017). The system of hypothalamic neurons, distinguishable both in terms of subnucleus and neurotransmitter/neurohormone profile, was specified in Maude. State-space analysis revealed activation configurations consistent with new experimental findings that contradicted current understandings. For example, high food-intake had been associated only with high levels of activity of arcuate nucleus neurons secreting Agouti-related peptide (AgRP neurons). Search of the model state-space found many configurations consistent with that relationship but also found some configurations consistent with new findings that *low* AgRP neuron activity also can be associated with high food intake. The model showed how overall activity patterns, comprising all of the neurons modeled, which were consistent with high food-intake differed between those having low versus high AgRP neuron activity. The differences stand as model predictions. A similar approach based on a model of the monoaminergic neurotransmitter system was used to study neuroadaptation to chronic antidepressant administration due to changes that are known to occur in various neurotransmitter/neurohormone receptor types. State-space analysis found many configurations adapted to chronic selective serotonin reuptake inhibitors (SSRIs) that nevertheless did not have serotonin levels that had been elevated to therapeutic levels. These findings provide an explanation for the low clinical efficacy of SSRIs, since elevated serotonin is thought to be the therapeutic mechanism of SSRIs. The analysis also uncovered chronic antidepressant combinations that should be more effective than single SSRIs in treating depression.

Maude models of some of the molecular interactions that underlie Alzheimer Disease (AD) have also appeared (Anastasio 2011, 2013a, 2014a, b, 2015). These studies were based on the amyloid hypothesis, which posits that AD results from buildup of the amyloid-beta peptide (Hardy and Selkoe 2002). A subset of the molecular and cellular interactions believed to underlie the regulation of amyloid-beta is diagrammed in Fig. 1.

Space constraints prohibit description of all of these interactions. They are based directly on findings from the primary literature and described in detail in Anastasio (2011). Each of 36 declarations specified how the level of biological activity of one model element, representing a molecular species, was determined through interaction among the others. Many model elements could assume multiple levels, so the entire state space was too large to be managed on a desktop computer. As an alternative to high-performance computing, the declarations, first specified as equations, were then divided into subsets that could be converted to rules. In this way, the space of configurations involving a subset of interactions could be analyzed using state-space search and temporal logical, while the rest of the interactions took place in the background.

An example of an equation in the Maude model of amyloid-beta regulation is



Application of Declarative Programming in Neurobiology, Fig. 1 Schematic diagram of many of the cellular and molecular interactions that regulate the level of the peptide amyloid-beta

```
ceq [PPARexpress] : PPARgene(G) NSAID
(X) cytokine(Y) PPAR(Z) =
PPARgene(G) NSAID(X) cytokine(Y) PPAR
(max(0, G + (X - Y)) * G)
if Z =/= max(0, G + (X - Y)) * G.
```

This equation, labeled `PPARexpress`, describes the interaction that determines the level of the peroxisome proliferator-activated receptor (PPAR). The operator `PPAR(Z)` assigns integer level `Z` to `PPAR`, and the equation describes how that level is determined by the presence of its gene PPARgene and the levels of cytokines and nonsteroidal anti-inflammatory drugs (NSAIDs). This equation is conditional (`ceq`) and executes only if doing so changes the level assigned by `PPAR`. To convert this conditional equation to a conditional rule, it would be relabeled `crl` and = would be replaced with =>.

The model depicted in Fig. 1 incorporates the hypothesis that mild cerebrovascular disease (CVD) can trigger amyloid-beta accumulation (Scheibel et al. 1989; de la Torre 2009). Simulations and analysis were focused on compounds known to modulate the activities of model elements. Specifically, cilnidipine blocks hypoxia-inducible factor (HIF) (Oda et al. 2009) while ifenprodil blocks caspase 3 (casp3) (Dave et al. 2003). It was of interest to see whether a HIF-blocker or a casp3-blocker would be more effective in reducing the rise of amyloid-beta in the presence of CVD in the model.

In the model, amyloid-beta rose from its normative level of 4 to the pathological level of 8 in the presence of CVD (Table 1). NSAIDs, known to increase PPAR levels in vitro (Sastre et al. 2006), and HIF-blocker each separately held the rise of amyloid-beta to 5 while casp3-blocker only held it to 7. NSAID and HIF-blocker together held amyloid-beta to its normative level of 4 but casp3-blocker provided no further benefit in combination with NSAID or HIF-blocker or both.

These simulations showed that an HIF-blocker was more effective than a casp3-blocker in reducing the rise of amyloid-beta in the face of CVD. We then used temporal-logic analysis to show why this effect occurs. Temporal-logic analysis provides answers to questions such as whether a particular property is always true, never true, or only true after some other property becomes true. For the AD model, temporal-logic analysis was used to determine the value of propositions of the form ~ `hasAPO U cytACT`, where `hasAPO` is the property that apoptosis has occurred, `cytACT` is the property that cytokines have been activated, and ~ and U are the temporal-logic operators "not" and "until," respectively. Maude demonstrated that this proposition is true, meaning that apoptosis will not occur until cytokines are activated in the model. That is important because apoptosis activated caspase 3 in the model. Similar analyses revealed that cytokines were not activated if HIF-blocker was present. Taken together, the temporal-logic analyses show that casp3-blocker was not effective in combination with HIF-blocker because HIF-blocker already prevented caspase 3 activation.

Application of Declarative Programming in Neurobiology, Table 1 Modeling the effects of nonsteroidal anti-inflammatory drugs (NSAIDs) and blockers of hypoxia-inducible factor and caspase 3 (HIF-blocker and casp-blocker) on the level of amyloid-beta in the presence of incipient cerebrovascular disease (CVD)

CVD	NSAID	HIF-blocker	Casp-blocker	Amyloid-beta
0	0	0	0	4
1	0	0	0	8
1	1	0	0	5
1	0	1	0	5
1	0	0	1	7
1	1	1	0	4
1	1	0	1	5
1	0	1	1	5
1	1	1	1	4

These analyses made several testable predictions:

1. An HIF-blocker will be effective in reducing amyloid-beta levels in the presence of incipient CVD.
2. A casp-blocker will be less effective than an HIF-blocker.

3. A casp-blocker will provide no further benefit if administered in conjunction with an HIF-blocker.
4. NSAIDs will be more effective in combination with an HIF-blocker than alone.

Conclusion

Declarative programming can be applied in the simulation and temporal-logic analysis of complex neurobiological processes, whether normal or pathological. The phenomenon at issue in the above example, the regulation of amyloid-beta, occurs predominantly on the molecular level. However, term-rewriting and declarative programming can also be used to create and analyze models of neurobiological phenomena at all scales, from molecular to cellular to network to whole brain, or to create multiscale models in order to understand phenomena arising from interactions across multiple scales.

References

Anastasio TJ (2011) Data-driven modeling of Alzheimer disease pathogenesis. J Theor Biol 290:60–72

Anastasio TJ (2013a) Exploring the contribution of estrogen to amyloid-beta regulation: a novel multifactorial computational modeling approach. Front Pharmacol 4:16

Anastasio TJ (2013b) Computational search for hypotheses concerning the endocannabinoid contribution to the extinction of fear conditioning. Front Comput Neurosci 7:74

Anastasio TJ (2014a) Computational identification of potential multitarget treatments for ameliorating the adverse effects of amyloid-beta on synaptic plasticity. Front Pharmacol 5:1

Anastasio TJ (2014b) Temporal-logic analysis of microglial phenotypic conversion with exposure to amyloid-β. Mol Bio Syst 11:434–453

Anastasio TJ (2015) Computational identification of potential multi-drug combinations for reduction of microglial inflammation in Alzheimer disease. Front Pharmacol 6:116

Camacho MB, Anastasio TJ (2017) Computational model of antidepressant response heterogeneity as multi-pathway neuroadaptation. Front Pharmacol 8:925

Cardone F, Hindley JR (2006) History of Lambda-calculus and combinatory logic. In: Gabbay DM, Woods J (eds) Handbook of the history of logic. Elsevier, Amsterdam

Clavel M, Durán R, Eker S, Lincoln P, Martí-Oliet N, Meseguer J, Talcott C (2007) All about Maude: a high-performance logical framework: how to specify, program, and verify systems in rewriting logic. Springer, Berlin

Dave JR, Williams AJ, Moffett JR, Koenig ML, Tortella FC (2003) Studies on neuronal apoptosis in primary forebrain cultures: neuroprotective/anti-apoptotic action of NR2B NMDA antagonists. Neurotox Res 5(4):255–264

de la Torre JC (2009) Cerebrovascular and cardiovascular pathology in Alzheimer's disease. Int Rev Neurobiol 84:35–48

Eker S, Knapp M, Laderoute K, Lincoln P, Meseguer J, Sonmez K (2002) Pathway logic: symbolic analysis of biological signaling. Pac Symp Biocomput 2002:400–412

Fisher J, Henzinger TA (2007) Executable cell biology. Nat Biotechnol 25(11):1239–1249

Hardy J, Selkoe DJ (2002) The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 297(5580): 353–356

Huth M, Ryan M (2004) Logic in computer science: modelling and reasoning about systems. Cambridge University Press, Cambridge, MA

Kain RY (1972) Automata theory: machines and languages. McGraw-Hill, New York

Meseguer J (2012) Twenty years of rewriting logic. J Log Alg Prog 81(7–8):721–781

Oda S, Oda T, Takabuchi S, Nishi K, Wakamatsu T, Tanaka T, Adachi T, Fukuda K, Nohara R, Hirota K (2009) The calcium channel blocker cilnidipine selectively suppresses hypoxia-inducible factor 1 activity in vascular cells. Eur J Pharmacol 606(1–3):130–136

Sastre M, Dewachter I, Rossner S, Bogdanovic N, Rosen E, Borghgraef P, Evert BO, Dumitrescu-Ozimek L, Thal DR, Landreth G, Walter J, Klockgether T, van Leuven F, Heneka MT (2006) Nonsteroidal anti-inflammatory drugs repress beta-secretase gene promoter activity by the activation of PPARgamma. Proc Natl Acad Sci USA 103(2):443–448

Scheibel AB, Duong TH, Jacobs R (1989) Alzheimer's disease as a capillary dementia. Ann Med 21(2):103–107

Tabe-Bordbar S, Anastasio TJ (2016) Computational analysis of the hypothalamic control of food intake. Front Comput Neurosci 10:27

Talcott C (2008) Pathway logic. In: Bernardo M, Degano P, Zavattaro G (eds) Lecture notes in computer science. Springer, Berlin, pp 21–53

Applications of Information Theory to Analysis of Neural Data

Simon R. Schultz[1], Robin A. A. Ince[2] and Stefano Panzeri[3,4]
[1]Department of Bioengineering, Imperial College London, London, UK
[2]School of Psychology, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK
[3]Center for Neuroscience and Cognitive Systems, Istituto Italiano di Tecnologia, Rovereto, Italy
[4]Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK

Definition

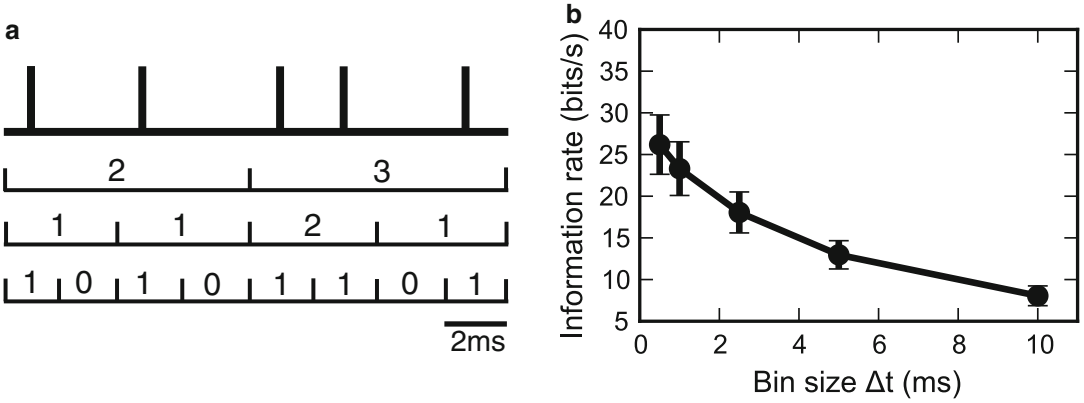
Information theory is a practical and theoretical framework developed for the study of communication over noisy channels. Its probabilistic basis and capacity to relate statistical structure to function make it ideally suited for studying information flow in the nervous system. It has a number of useful properties: it is a general measure sensitive to any relationship, not only linear effects; it has meaningful units which in many cases allow direct comparison between different experiments; and it can be used to study how much information can be gained by observing neural responses in single trials, rather than in averages over multiple trials. A variety of information-theoretic quantities are commonly used in neuroscience – (see entry ▶ "Summary of Information-Theoretic Quantities"). In this entry we review some applications of information theory in neuroscience to study encoding of information in both single neurons and neuronal populations.

Detailed Description

Information Analysis of Spike Trains to Investigate the Role of Spike Times in Sensory Coding

Mutual information is a widely used tool to study how spike trains encode sensory variables. A typical application of mutual information to spike train analysis is to use it to compare the information content of different representations of neural responses that can be extracted from spike trains. The neural code used by a neuron is often defined operationally as the smallest set of response variables that carries all (or almost all) the information contained in the spike train of the neuron. Mutual information is used to quantify the information content of increasingly complex representations of the neural response, and the simplest representation that carries the most information is chosen as the putative neural code.

An example of this general approach is the investigation of the role of spike times in encoding information. The most established hypothesis on how sensory information is represented in the brain is the spike count-coding hypothesis (Adrian 1928) which suggests that neurons represent information by the number of spikes discharged over some relevant time window. Another hypothesis is the spike timing encoding hypothesis, which suggests that the timing of spikes may add important information to that already carried by spike counts (Rieke et al. 1997; Panzeri et al. 2001). Information theory can be used to understand the role of spike times in carrying sensory information, by using it to characterize the temporal resolution needed to read out the information carried by spike trains. This can be performed by sampling the spike train at different temporal precisions, Δt, (Fig. 1a) and computing the information parametrically as a function of Δt (Ruyter et al. 1997). The temporal precision required to read the temporal code then can be defined as the largest Δt that still provides the full information obtained at higher resolutions. If this precision is equal to the overall length of the window over which neurons carry information, then information is carried only by the number of spikes. As an example, we carried out this type of analysis on the responses of neurons from the VPm thalamic nucleus of rats whose whiskers were stimulated by fast white noise deflections (Montemurro et al. 2007). We found that the temporal precision Δt at which neurons transmitted information about whisker deflections was finer than 1 ms (Fig. 1b), suggesting that these neurons



Applications of Information Theory to Analysis of Neural Data, Fig. 1 *Effect of temporal resolution of spike times on information*. (**a**) The response of a neuron is initially recorded as a series of spike times. To investigate the temporal resolution at which spike times carry information, the spike train is binned at a variety of different time resolutions, by labeling the response at each time with the number of spikes occurring within that bin, thereby transforming the response into a discrete integer sequence. (**b**) The information rate (information per unit time) about whisker deflections carried by VPm thalamic neurons as a function of bin width, Δt, used to bin neural responses. Information rate increased with bin resolution up to 0.5 ms, the limit of the experimental setup. This shows that a very fine temporal resolution is needed to read out the sensory messages carried by these thalamic spike trains. (Figure reprinted with permission from Ince et al. 2010)

use high-precision spike timing to carry information.

Information Analysis of Local Field Potentials to Examine the Information Content of Network Oscillations

Information analysis in neuroscience is not limited only to spike train analysis, but it has been used also to study the measure of massed population activity, such as local field potentials (LFPs) (Buzsáki et al. 2012). LFPs are operationally defined as the low-pass filtered extracellular potential measured by an extracellular intracranial electrode. There are at least three reasons why LFPs are widely used in neuroscience. The first is that they are more easily and stably recorded in chronic settings than is the spiking activity of individual neurons. The second is that the LFP captures key integrative synaptic processes and aspects of subthreshold neural activity that cannot be measured by observing the spiking activity of a few neurons alone (Einevoll et al. 2013). The third is that LFPs are more sensitive to network oscillations than measures of spiking activity from small populations. LFPs from a sensory area typically show a power spectrum containing fluctuations over a wide range of frequencies, from <1 to 100 Hz or so. Given that the power of oscillatory activity typically increases during the presentation of a sensory stimulus, many authors have speculated that this oscillatory activity plays a role in brain communication and in particular in sensory-related computations. However, understanding the function of these oscillations has remained elusive and controversial. To gain insights into the function of oscillations in sensory encoding, it is important to understand how they contribute to the representation of the natural sensory environment.

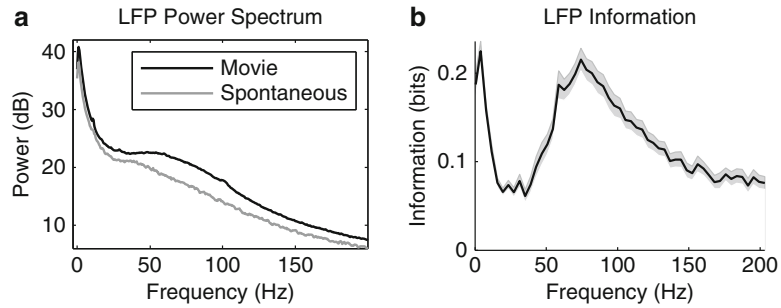
This problem can be addressed by quantifying the oscillation power in any given trial in response to different stimuli and then computing the information gained by the power at each frequency. Since the power is a continuous variable, the computation of its information is potentially more difficult than the one based on discrete variables like the spike train ones described above. There are at least two ways to solve this problem. The first is to discretize the power in a number of equi-populated bins and to use bias corrections to eliminate the bias. The second is to fit the data to a parametric distribution. In this case, it is worth

reminding that the power computed with most spectral methods follows a chi-square distribution, and thus, its square or third can reasonably be well approximated by a Gaussian distribution (Magri et al. 2009). This makes the computation of information relatively straightforward. The third potential approach is to use binless methods such as Nearest Neighbors approaches (Kraskov et al. 2004). We tried out these methods extensively on computation of information in power of LFPs, obtaining very similar results with all approaches (see, e.g., Magri et al. 2009).

We applied this method to recordings from primary visual cortex of anesthetized macaques during stimulation with naturalistic color movies (Belitski et al. 2008; Magri et al. 2012a). This revealed, for the first time, how information about the naturalistic sensory environment is spread over the wide range of frequencies expressed by cortical activity. Although the broadband nature of the spectrum suggests a contribution to coding from many frequency regions, we found that only two separate frequency regions contribute to coding: the low-frequency range and the gamma range (Belitski et al. 2008; see Fig. 2 below). Interestingly, low- and high-frequency ranges act as perfectly complementary or "orthogonal" information channels: they share neither signal (i.e., stimulus information) nor "noise" (i.e., trial to trial variability for a fixed stimulus). This finding has several implications. First, it shows that, despite the broadband spectrum, only a small number of privileged frequency scales are involved in stimulus coding. Second, it suggests that high-frequency and low-frequency oscillations are generated by different stimulus-processing neural pathways. Third, the finding that different frequency bands code different sensory features in separate, truly independent information channels reinforces the concept of "cortical multiplexing" that we proposed above.

Information Analysis of Imaging Data to Study Neural Population Coding or Coupling Between Different Neural Signals

The analysis tools that we have described above have, to date, largely been applied to spike train and time series data recorded using electrophysiological techniques. However, in recent years, imaging technologies have been developed which are capable of resolving neural signaling at systems' cellular and subcellular resolutions on a single trial basis (Denk et al. 1990, 1994; Stosiek et al. 2003; Chen et al. 2013). One way to apply information-theoretic tools to the analysis of such imaging data is to convert the data to a "spike train," for instance, by applying an algorithm for the detection of action potential-evoked calcium transients to calcium imaging data (Oñativia et al. 2013). Such an approach has been used to perform information-theoretic analysis of simultaneously recorded populations of cerebellar Purkinje cell complex spikes extracted from in vivo calcium imaging movies (Schultz et al. 2009). However,



Applications of Information Theory to Analysis of Neural Data, Fig. 2 *The visual information carried by LFP power at different frequencies*. (**a**) LFP power spectrum of V1 recordings anesthetized macaques either during spontaneous activity in the *dark* (*dashed line*) or during the presentation of a color movie stimulus (*solid line*). (**b**) Information about the movie stimulus carried by LFP power at different frequencies. The area indicates the SEM. (Reproduced from Magri et al. 2012a)

the use of imaging data may also allow a wider set of questions to be approached than can be examined electrophysiologically, by directly examining patterns of pixel intensities.

Another interesting application of information theory to neuroimaging data regards its use for understanding the nature of the coupling between neural activity and fMRI responses. In fact, although there is evidence that fMRI BOLD responses reflect neural activity, it is not clear whether the BOLD signal reflects only the total power of massed neural activity, or only the power in a given band, or rather the relationships between powers of neural activity in different frequency bands. This problem can be cast theoretically into quantifying whether more information about BOLD can be gained from simultaneously observing the power of neural activity in two or more bands of neural activity than the information gained by observing either band alone. Because mutual information captures all the ways a signal may statistically relate to another, finding that another signal carries extra information demonstrates that this signal truly provides some information that cannot be possibly obtained from the first one. This does not necessarily hold when using methods that capture only specific relationships between signals. For example, an increase in predictability based on linear models may reflect both additional information from the second regressor as well as information that was already present in the first regressor but was not captured by the linear assumption. Application of this idea to simultaneous recording of LFPs and fMRI BOLD in primary visual cortex showed that the beta and alpha bands carry information about BOLD that complements that carried by the gamma band, the band that most correlates to the BOLD signal (Magri et al. 2012b).

Since imaging signals such as fMRI have an analogue rather than discrete nature, the practicality of application of information theory to analogue brain signals is crucially dependent upon the development of appropriate regularization and dimensionality reduction algorithms. These might stem from simple yet efficient discretization algorithms (Belitski et al. 2008; Magri et al. 2009), Nearest Neighbors regularization algorithms (Kraskov et al. 2004), the use of manifold learning techniques for nonlinear dimensionality reduction (Roweis and Saul 2000; Seung and Lee 2000; Gan 2006), and/or the evaluation of information through a decoding step (Quian Quiroga and Panzeri 2009).

Acknowledgments Research is supported by the SI-CODE (FET-Open, FP7-284533) project and by the ABC and NETT (People Programme Marie Curie Actions PITN-GA-2011-290011 and PITN-GA-2011-289146) projects of the European Union's Seventh Framework Programme FP7 2007–2013.

References

Adrian ED (1928) The basis of sensation. Norton, New York. http://psycnet.apa.org/psycinfo/1928-01753-000. Accessed 17 Jan 2014

Belitski A, Gretton A, Magri C, Marayama Y, Montemurro MA, Logothetis NK, Panzeri S (2008) Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. J Neurosci 28: 5696–5709

Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents – EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407–420

Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499:295–300

Denk W, Strickler JH, Webb WW (1990) Two-photon laser scanning fluorescence microscopy. Science 248:73–76

Denk W, Delaney KR, Gelperin A, Kleinfeld D, Strowbridge BW, Tank DW, Yuste R (1994) Anatomical and functional imaging of neurons using 2-photon laser scanning microscopy. J Neurosci Methods 54: 151–162

Einevoll GT, Kayser C, Logothetis NK, Panzeri S (2013) Modelling and analysis of local field potentials for studying the function of cortical circuits. Nat Rev Neurosci 14:770–785

Gan JQ (2006) Feature dimensionality reduction by manifold learning in brain-computer interface design. In: Proceedings of the 3rd international workshop on brain-computer interfaces, Graz, pp 28–29. http://cswww.essex.ac.uk/Research/BCIs/BCI06_GAN1.pdf. Accessed 17 Jan 2014

Ince RAA, Mazzoni A, Petersen RS, Panzeri S (2010) Open source tools for the information theoretic analysis of neural data. Front Neurosci 4:62–70

Kraskov A, Stögbauer H, Grassberger P (2004) Estimating mutual information. Phys Rev E 69:66138

Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S (2009) A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. BMC Neurosci 10:81
Magri C, Mazzoni A, Logothetis NK, Panzeri S (2012a) Optimal band separation of extracellular field potentials. J Neurosci Methods 210:66–78
Magri C, Schridde U, Murayama Y, Panzeri S, Logothetis NK (2012b) The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. J Neurosci 32:1395–1407
Montemurro MA, Panzeri S, Maravall M, Alenda A, Bale MR, Brambilla M, Petersen RS (2007) Role of precise spike timing in coding of dynamic vibrissa stimuli in somatosensory thalamus. J Neurophysiol 98: 1871–1882
Oñativia J, Schultz SR, Dragotti PL (2013) A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. J Neural Eng 10:046017
Panzeri S, Petersen RS, Schultz SR, Lebedev M, Diamond ME (2001) The role of spike timing in the coding of stimulus location in rat somatosensory cortex. Neuron 29:769–777
Quian Quiroga R, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci 10: 173–185
Rieke F, Bialek W, Warland D, de Ruyter van Steveninck RR (1997) Spikes: exploring the neural code. Bradford Book
Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290: 2323–2326
Ruyter D, van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W (1997) Reproducibility and variability in neural spike trains. Science 275: 1805–1808
Schultz SR, Kitamura K, Post-Uiterweer A, Krupic J, Häusser M (2009) Spatial pattern coding of sensory information by climbing fiber-evoked calcium signals in networks of neighboring cerebellar purkinje cells. J Neurosci 29:8005–8015
Seung HS, Lee DD (2000) The manifold ways of perception. Science 290:2268–2269
Stosiek C, Garaschuk O, Holthoff K, Konnerth A (2003) In vivo two-photon calcium imaging of neuronal networks. Proc Natl Acad Sci U S A 100:7319–7324

Artificial Retina

▶ Vision Prosthesis
▶ Visual Prosthesis, Epiretinal Devices
▶ Visual Prosthesis, Subretinal Devices

Artificial Silicon Retina (ASR)

▶ Visual Prosthesis, Subretinal Devices

Artificial Vision

▶ Prosthetic Vision, Perceptual Effects
▶ Vision Prosthesis
▶ Visual Prosthesis, Epiretinal Devices

Associations and Rewards in the Auditory Cortex

Michael Brosch
Leibniz Institute for Neurobiology, Magdeburg, Germany

Definition

In psychology, the term *association* refers to a connection between different elementary mental entities (sensations, thoughts, feelings; Dudai 2002). Aside from innate, reflex-like associations, novel associations are typically acquired during learning. In Pavlovian/classical conditioning, a stimulus-stimulus association is formed by repetitively pairing an initially neutral stimulus with a biologically significant unconditioned stimulus that automatically triggers an unconditioned behavioral response. In instrumental/operant conditioning, a stimulus-response association is formed in the presence of reinforcers. Reinforcers can be either positive, such as water, food, money, and brain stimulation reward, and result in an increase in the probability of a response to the stimulus. Negative reinforcers (e.g., footshocks, airpuffs, money loss) decrease the probability of a response to the stimulus. If the reinforcer is removed, the learned associations risk extinction.

Detailed Description

Classical View of Brain Structures Reflecting Associations and Reinforcement

Traditionally associative functions are assigned to the so-called association cortex (Creutzfeldt 1983). Inspired by "associationism" (the philosophical doctrine that the mind learns and construes the world bottom up by associating mental entities) and based on anatomical considerations, Flechsig originally defined the association cortex as that part of the cerebral cortex that appeared to lack direct afferences from the senses and efferences to peripheral motor structures. He proposed that the association cortex provides the substrate for the fusion of primary sensations to obtain ideas of objects as a whole. This idea is challenged, inter alia, by studies demonstrating associative functions already in primary sensory cortices, which have direct afferences from the senses. In addition, the sensory cortex is affected by reinforcement (Shuler and Bear 2006; Brosch et al. 2011a; Arsenault et al. 2013; Weis et al. 2013), which generally is thought to involve the limbic system, including the hypothalamus, amygdala, hippocampus, septal nuclei, ventral tegmental area, and anterior cingulate gyrus. These findings put into question the existence of unisensory cortical areas at all (Ghazanfar and Schroeder 2006).

Learning

Learning associations between stimuli, between stimuli and behavioral responses, or between stimuli and reinforcers change the auditory cortex, such as the feature sensitivity (spectrotemporal receptive field) of neurons in the auditory cortex, their response strength and response latency, as well as interneuronal synchrony (Scheich and Brosch 2013; Shepard et al. 2013). In detection tasks, the direction of the receptive field change at the frequency of the conditioned tone depends on the valence of the tone (Scheich et al. 2011); if it is negative (associated with punishment), the response to the tone is increased (receptive field is sensitized at the conditioned tone frequency); if it is positive (associated with reward), the response is decreased (receptive field is suppressed at the conditioned tone frequency). In frequency discrimination tasks, slopes of spectral tuning curves become sharper around the conditioned frequencies. In categorization tasks, stimuli of the same category evoke (spatiotemporal) neuronal activity patterns that are more similar to each other than those evoked by stimuli of other categories.

These changes coincide with or may even form the basis of changes in representation maps (e.g., tonotopic frequency map) in the auditory cortex that occur, at least transiently, after learning (Shepard et al. 2013; Grosso et al. 2015). They may also underlie plasticity of neuronal mass activity, as revealed by electro- and magnetoencephalography (EEG, MEG) or functional magnetic resonance imaging (fMRI) (Rüsseler et al. 2005).

Neuromodulators

The formation of novel associations in the auditory cortex requires the involvement of neuromodulators (Shepard et al. 2013). Some of the described changes can also be mimicked by repetitively pairing auditory stimuli with electrical stimulation of neuromodulatory systems, such as the ventral tegmental area (dopamine; Huang et al. 2016a), the nucleus basalis of Meynert (acetylcholine; Bakin and Weinberger 1996), the locus coeruleus (norepinephrine; Martins and Froemke 2015), or the vagus nerve (triggering widespread release of neuromodulators; Engineer et al. 2011). Formation of associations in the auditory cortex also involves cognitive factors: changes in the auditory cortex and learning do not occur when an animal is passively exposed to the stimulus-reward pairings of another animal while it was instrumentally conditioned (Blake et al. 2006).

Neuronal Correlates of Association

Sustained firing and slow firing changes may provide neuronal correlates of associations between mental entities (Brosch et al. 2011a). The main condition necessary for the emergence of slow firing changes is that subjects have learnt that conditioned associations are contingent on reinforcers. When the reinforcer is removed, the slow

firing changes disappear within a few trials, concomitantly with behavioral changes (Huang et al. 2016b). These firings may be related to the contingent negative variation (Walter et al. 1964), an event-related potential that can be obtained in electroencephalography in humans. It is considered to reflect the contingency and the contiguity of two events that have a motivational "value."

Different types of events may be associated through event-related slow firing changes (Brosch et al. 2011a). Events that trigger such changes can be auditory (and possibly visual) stimuli, reinforcers, and behavioral responses, like grasping. An event with which slow firing changes end is the reinforcer. Slow firing changes have been observed between (1) an auditory stimulus and a reinforcer, (2) a behavioral response and an auditory stimulus, and (3) a behavioral response and another behavioral response.

The association between behaviorally significant events provided by slow firing changes might even be directed in some cases, that is, this type of firing might provide some sort of either prospective coding of an upcoming event or retrospective coding about a preceding event.

Neuronal Correlates of Rewards

In animals actively engaged in listening to auditory stimuli, direct reflections of rewards have been demonstrated in the auditory cortex. Neuronal activity in the auditory cortex varied with the size of the delivered reward and the size of the reward that was expected to be earned in a future auditory task, as well as the magnitude of the mismatch between the expected and delivered reward (the reward prediction error) (Brosch et al. 2011b). Reflections of rewards are also present in other sensory cortices (Shuler and Bear 2006; Arsenault et al. 2013).

References

Arsenault J, Nelissen K, Jarraya B, Vanduffel W (2013) Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. Neuron 77(6):1174–1186

Bakin JS, Weinberger NM (1996) Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. Proc Natl Acad Sci USA 93:11219–11224

Blake DT, Heiser MA, Caywood M, Merzenich MM (2006) Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. Neuron 52:371–381

Brosch M, Selezneva E, Scheich H (2011a) Formation of associations in auditory cortex by slow changes of tonic firing. Hear Res 271:66–73

Brosch M, Selezneva E, Scheich H (2011b) Representation of reward feedback in primate auditory cortex. Front Syst Neurosci 5:5

Creutzfeldt OD (1983) Cortex Cerebri: Leistung, Strukturelle und Funktionelle Organisation der Hirnrinde. Springer, Berlin/Heidelberg/New York

Dudai Y (2002) Memory from a to Z. Oxford University Press, Oxford

Engineer ND, Riley JR, Seale JD, Vrana WA, Shetake JA, Sudanagunta SP, Borland MS, Kilgard MP (2011) Reversing pathological neural activity using targeted plasticity. Nature 470:101–104

Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? Trends Cogn Sci 10:278–285

Grosso A, Cambiaghi M, Concina G, Sacco T, Sacchetti B (2015) Auditory cortex involvement in emotional learning and memory. Neuroscience 299:45–55

Huang Y, Mylius J, Scheich H, Brosch M (2016a) Tonic effects of the dopaminergic ventral midbrain on the auditory cortex of awake macaque monkeys. Brain Struct Funct 221:969–967

Huang Y, Matysiak A, König R, Heil P, Brosch M (2016b) Persistent neural activity in auditory cortex is related to auditory working memory in humans and nonhuman primates. eLife 5. Pii: e15441. https://doi.org/10.7554/eLife

Martins AR, Froemke RC (2015) Coordinated forms of noradrenergic plasticity in the locus coeruleus and primary auditory cortex. Nat Neurosci 18:1483–1492

Rüsseler J, Nager W, Möbes J, Münte TF (2005) Cognitive adaptations and neuroplasticity: lessons from event-related brain potentials. In: König R, Heil P, Budinger E, Scheich H (eds) Auditory cortex: towards a synthesis of human and animal research. Lawrence Erlbaum, Mahwah, pp 467–484

Scheich H, Brosch M (2013) Task-related activation of the auditory cortex. In: Cohen YE, Popper AN, Fay RR (eds) Neural correlates of auditory cognition, springer handbook of auditory research. Springer, New York/Heidelberg/Dordrecht/London

Scheich H, Brechmann A, Brosch M, Budinger E, Ohl FW, Selezneva E, Stark H, Tischmeyer W, Wetzel W (2011) Behavioral semantics of learning and crossmodal processing in auditory cortex: the semantic processor concept. Hear Res 271:3–15

Shepard KN, Kilgard MP, Liu RC (2013) Experience-dependent plasticity and the auditory cortex. In: Cohen YE, Popper AN, Fay RR (eds) Neural correlates of auditory cognition, springer handbook of auditory

research. Springer, New York/Heidelberg/Dordrecht/London
Shuler MG, Bear MF (2006) Reward timing in the primary visual cortex. Science 311:1606–1609
Walter WG, Cooper R, Aldridge VJ, McCallum WC, Winter AL (1964) Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. Nature 203:380–383
Weis T, Puschmann S, Brechmann A, Thiel CM (2013) Positive and negative reinforcement activate human auditory cortex. Front Hum Neurosci 7:842

Associative Learning

► Eye-Blink Conditioning

Attentional Top-Down Modulation, Models of

Philipp Schwedhelm[1,3] and Stefan Treue[1,2,3]
[1]Cognitive Neuroscience Laboratory, German Primate Center, Göttingen, Germany
[2]Faculty of Biology and Psychology, Göttingen University, Göttingen, Germany
[3]Bernstein Center for Computational Neuroscience, Göttingen University, Göttingen, Germany

Definition

Attention – the ability of a sensory system to facilitate the processing of specific information at the expense of disregarding the remainder.
Bottom-up processes – information processing in the nervous system that operates in a feedforward way, advancing from sensory organs or areas at a low level of the cortical processing hierarchy.
Top-down influence – modulatory signals in the nervous system that originate from areas at a high level of the cortical processing hierarchy, influencing information processing in lower areas.
Saliency – a measure of the magnitude of the difference of a stimulus from its neighbors in space and time.

Detailed Description

The Case for Attention

Evolution has provided humans and other highly evolved species with powerful sensory systems. While our cortical processing capacity has also evolved and grown impressively, the torrent of information provided by our sensors far outstrips our ability to process it all. In addition, most of the sensory information picked up at any moment has little importance for our survival. Complex nervous systems faced with this challenge have developed sophisticated selection mechanism to identify the most relevant incoming information and to focus processing resources (and ultimately perception) onto this small fraction. This process is called attention and for the purpose of this entry can be defined as the selective modulation of sensory information based on its assumed behavioral relevance.

Bottom-Up Versus Top-Down

The selection processes underlying attention need to fulfill two requirements: on the one hand their ubiquitous (central and incessant) role in the continuous stream of perceptual decisions requires that they operate efficiently and as fast as possible. At the same time, the selection processes' purpose of dynamically identifying the most relevant components of the sensory input demands harnessing as much of the cognitive power of the species' central nervous system as possible.

These seemingly incompatible demands, efficient and fast vs. computationally demanding and thus slow, have created two flavors of selection:

1. A bottom-up (automatic, exogenous) attentional selection that exploits the realization that the most informative aspects of our sensory environments are those where one stimulus differs from their neighbors in space and time. This local saliency can be identified and enhanced by simple feedforward filter mechanisms embedded throughout the processing of sensory signals in the nervous system.

2. A top-down (voluntary, endogenous) attentional selection that integrates any information available to the organism about the current situation to make the most informed decision about which sensory input component represents the most relevant information in the given situation.

In the visual domain, this distinction is well illustrated with visual search tasks: If we are confronted with a fairly homogenous visual scene, any outlier will be identified, enhanced, and selected by the continuous parallel computation of local saliency, creating the perceptual "pop-out" characteristic of simple search tasks where the features of the target stimulus differ substantially from the distribution of features of the distractors. Conversely, a target stimulus, which is less distinct, either because it is defined as a conjunction of more than one feature or because it does not differ substantially from the distribution of distractor features, does not pop out, but rather requires a more demanding and correspondingly slower selection process.

Taking a Computational Approach to Attention

Here we illustrate how the attentional modulation of sensory information processing is implemented in computational models. Due to the brevity of the entry, we focus on a few examples of models of top-down attentional modulation in the visual system of man and other primates.

One of the most influential computational models of visual attention is the **feature integration theory** (FIT; Treisman and Gelade 1980). In the FIT, information about different features of stimulus, such as its shape, color, orientation, and movement, is extracted in parallel, automatically and effortlessly through a system of feature maps, which topographically represent the spatial distribution of specific features in the visual scene. This process detects and locates a target stimulus defined by a single unique feature value (such as the color red) because it is represented by a unique hotspot in a single feature map (with each distractor represented by a hotspot in its corresponding feature map, such as the one for the color blue). This target detection is very quick and is unaffected by the numerosity of distractor stimuli, matching the experimental observation that human reaction times in such simple search tasks are independent of the number of distractor items. If the target stimulus is not defined by a single feature alone, but by a conjunction of multiple features, information from different feature maps needs to be integrated to detect and localize a target. This requires a serial process that actively integrates information from different maps to detect the target's unique feature conjunction at one topographical location, matching the linear increase in reaction time observed with an increase in the number of distracters in a conjunctive search task. The FIT proposes that this serial integration process is accomplished by means of a top-down, spatial "spotlight" of attention.

An alternative account for the pattern of reaction times in search experiments is offered by the **guided search theory** (GST, Wolfe 1994), which does not assume an attentional spotlight. Instead, the top-down attentional signal changes the weight of activation maps before they are combined to create a ranking of all present stimuli based on their likelihood to represent a target. The selection of stimuli is then again performed serially, from high to low probability, until the target stimulus is detected.

While the FIT and the GST emphasize the role of feature maps in attentional selection, the **theory of visual attention** (TVA; Bundesen 1990) takes a different approach. Here the selection of stimuli is dependent on their processing speed. Before a stimulus can be encoded in visual short-term memory and thus enter awareness, it needs to compete in a computational race with other stimuli. In the TVA top-down attention speeds up the processing of certain items, making them likely to win the race.

While the FIT, GST, and TVA have been developed to account for the perceptual data available at the time, more recent models of attention have been developed to capture data from single-cell recordings from monkey visual cortex. Two early conceptual models attempted to account for the enhanced neuronal response to attended stimuli

and the reduced response to unattended stimuli. The **biased competition model** of attention (Desimone and Duncan 1995) envisages a competition between the stimulus representation of attended and unattended stimuli that can be biased by a top-down attentional signal in favor of the attended stimulus' representation. The **feature similarity gain model** of attention (Treue and Martinez-Trujillo 1999) alternatively proposes that the enhancement of neural responses by attention reflects a process where top-down attentional signals enhance the gain of those neurons whose preferred features match the current attentional state of the organism, independent of the stimulus that currently activates a neuron.

These two conceptual models have inspired a large number of computational models. The most prominent of those are models that emphasize an interaction of top-down attention with the normalization process that creates the sigmoidal contrast response functions typical for neurons throughout sensory cortex. Multiple varieties of such **normalization models** of attention have been proposed (Ghose and Maunsell 2008; Boynton 2009; Ghose 2009; Lee and Maunsell 2009, 2010; Reynolds and Heeger 2009). They all emphasize the similarity, in perception, as well as in the neural encoding and also in the central role of the response normalization process between two influences on the strength a neural stimulus representation. One is the physical (bottom-up) strength of the stimulus (most directly represented by its contrast) and the other is the attentional weight (implemented as a kind of sensory prior) assigned to them through a top-down attentional signal.

Beyond models that emphasize response normalization, there have been numerous other approaches to model the attentional modulation of sensory information processing. They include the **selective tuning model** (Tsotsos et al. 2005) that proposes a layered network architecture (representing the hierarchy of cortical areas) to implement a spatial "spotlight of attention" that endows certain regions of the visual scene with enhanced processing. The **spiking network model** (Deco and Rolls 2005; Deco and Thiele 2011) places much more emphasis than any of the models discussed above on building its approach on biological components, such as spiking neurons and specific neurotransmitters.

A

The Integrated Saliency Map

It should be noted that almost all models of attention incorporate the concept of an **integrated saliency map** (Treue 2003), that is, a topographic representation of the stimuli in the current visual scene that combines their relative physical strength and their assumed behavioral relevance. This combination implements a weighing of bottom-up and top-down aspects of a stimulus, providing processing resources to strong unattended stimuli as well as to weak attended ones. While such an integrated saliency map. Is consistent with a number of perceptual phenomena and is ideally suited to guide eye movements across a visual scene, it is a matter of some debate which of the many topographically organized areas in the visual cortex represents this map or whether multiple such maps exist.

Similarly, while functional imaging and single-cell recording studies have implicated a network of frontoparietal areas in the guidance process that is necessary to appropriately allocate processing resources (Kastner and Ungerleider 2001; Corbetta and Shulman 2002), such anatomic specificity is rarely included in current computational models of attention.

Conclusion

In conclusion, in the last decade, a large number of computational models of top-down attention have been developed that can account for a large variety of perceptual and physiological aspects of the attentional modulation of sensory information processing. These models emphasize several core issues, such as the response normalization in cortical networks, the multistage nature of cortical information processing, and the concept of an integrated saliency map. Despite this progress much more work is needed to achieve a complete computational description of top-down attentional modulation.

Cross-References

- ▶ Hierarchical Models of the Visual System
- ▶ Working Memory, Models of

References

Boynton GM (2009) A framework for describing the effects of attention on visual responses. Vis Res 49: 1129–1143

Bundesen C (1990) A theory of visual attention. Psychol Rev 97:523–547

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

Deco G, Rolls ET (2005) Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. J Neurophysiol 94: 295–313

Deco G, Thiele A (2011) Cholinergic control of cortical network interactions enables feedback-mediated attentional modulation. Eur J Neurosci 34:146–157

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18(1): 193–222

Ghose GM (2009) Attentional modulation of visual responses by flexible input gain. J Neurophysiol 101: 2089–2106

Ghose GM, Maunsell JH (2008) Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. J Neurosci 28:5115–5126

Kastner S, Ungerleider LG (2001) The neural basis of biased competition in human visual cortex. Neuropsychologia 39:1263–1276

Lee J, Maunsell JH (2009) A normalization model of attentional modulation of single unit responses. PLoS One 4:e4651

Lee J, Maunsell JH (2010) Attentional modulation of MT neurons with single or multiple stimuli in their receptive fields. J Neurosci 30:3058–3066

Reynolds J, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185

Treisman A, Gelade G (1980) A feature-integration theory of attention. Cog Psychol 12:97–136

Treue S (2003) Visual attention: the where, what, how and why of saliency. Curr Opin Neurobiol 13: 428–432

Treue S, Trujillo JCM (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399(6736):575–579

Tsotsos JK, Liu Y, Martinez-Trujillo JC, Pomplun M, Simine E, Zhou K (2005) Attending to visual motion. Compu Vis Image Underst 100:3–40

Wolfe JM (1994) Guided search 2.0: a revised model of visual search. Psychon Bull Rev 1:202–238

Attractor Neural Network

- ▶ Hopfield Network

Auditory Brainstem Responses

Carles Escera[1,2,3,4] and Natàlia Gorina-Careta[1,2,3]

[1]Brainlab-Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain

[2]Institute of Neurosciences, University of Barcelona, Barcelona, Spain

[3]Institut de Recerca Sant Joan de Déu (IRSJD), Barcelona, Spain

[4]Institute for Brain, Cognition and Behavior (IR3C), University of Barcelona, Barcelona, Spain

Synonyms

Auditory evoked brainstem responses; Brainstem auditory evoked potentials (BAEP)

Definition

Auditory brainstem responses (ABRs) are the earliest auditory event-related potentials (cf. ▶ "Auditory Event-Related Potentials") elicited during the first 10 ms after stimulus onset in the anatomical relays of the auditory pathway. ABRs are typically recorded with scalp electrodes attached to the vertex and referenced to the earlobe or the mastoid. Two types of ABRs have been described: the transient-evoked responses and the sustained frequency-following response (FFR). While the transient-evoked ABRs are elicited to high-intensity clicks or tone bursts, FFRs can only be recorded to periodic auditory stimuli, and their duration extends to the length of the eliciting sound. In the ABRs elicited to periodic and complex auditory stimuli,

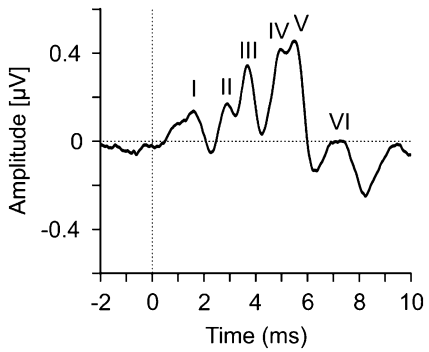
such as speech sounds or music, the two types of ABRs can be observed.

Detailed Description

Transient-Evoked ABRs

The first systematic observation of the human transient ABRs was by Jewett, Romano, and Williston in 1970 (Jewett et al. 1970; Jewett and Williston 1971) who described a series of six or seven fully visible waveforms spanning the first 10 ms after sound onset. ABRs are characterized for being fast, vertex-positive response peaks spaced at intervals of about 1 ms and are evoked by brief stimulus features, such as broadband clicks, tone bursts, or the abrupt onset of a sound (Pratt 2012).

The different ABR waveforms define a very characteristic morphology and are labelled in sequence by Roman numerals from I to VI (Fig. 1). Their amplitude is in the order of tenths of a microvolt, smaller than 0.5 mV (Pratt 2012), with the most positive component being wave V (Picton 2011; Pratt 2012). The first component of the sequence, wave I, is the most positive peak occurring at approximately 1 ms after stimulus onset and corresponds to the N1 wave of the electrocochleogram. Between waves I and V, the most prominent deflection is wave III, which occurs at a latency of about 3.5 ms. Wave II is a small peak halfway between wave I and III, and waves IV and VI can be observed in both sides of wave V. Wave IV is not always identified in humans, and when present it is often merged partially with wave V, yielding a wave IV-V complex (Picton 2011; Pratt 2012).



Auditory Brainstem Responses, Fig. 1 Morphology of the human transient ABR. The figure depicts the typical morphology of a healthy human adult ABR. It was obtained by averaging the responses to 6000 presentations of a click delivered bilaterally at 85 dB SPL at a rate of 19.3/second. The recording is between electrodes placed at FPz and the right mastoid. The EEG was acquired at 20 kHz sampling rate using a band-pass filter from 100 to 3000 Hz. Waves I to VI are identified at their typical latencies

The different ABR components are generated by distinct anatomic structures of the auditory pathway. The first components of the sequence, waves I and II, are generated exclusively by the auditory nerve (Gelfand 2010). More precisely, wave I originates in the ipsilateral distal portion of the auditory nerve in the region of the ganglion cells and wave II from the ipsilateral proximal portion of the auditory nerve in the vicinity of its entry into the brainstem (Stone et al. 2009; Pratt 2012). The later components, waves III, IV, and V, have contributions from more than one anatomical structure. Wave III has been attributed to the region of the ipsilateral cochlear nucleus and the superior olivary complex (Stone et al. 2009; Picton 2011). The generators of wave IV are still under debate, but all evidence points to bilateral multiple brainstem sources (Stone et al. 2009), between the superior olivary complex through the lateral lemniscus, with a possible contribution from the inferior colliculus (Pratt 2012). Wave V is originated in the contralateral distal lateral lemniscus and the inferior colliculus (Stone et al. 2009), so overall, the wave IV-V complex is attributed to bilateral generators in the region of the midbrain tegmentum (Starr et al. 2008). Wave VI has been attributed to the medial geniculate body (Gelfand 2010). Although these sources are the most accepted ones across the literature, there are still divergent opinions regarding the exact origins of the components from wave II onwards, as it is known that multiple generators in the auditory brainstem may be contributing to them (Gelfand 2010; Picton 2011).

The transient ABRs are typically elicited by broadband clicks or chirps (Fobel and Dau 2004; Maloff and Hood 2014) presented at a rate ranging from 10 to 20 stimuli per second. They are optimally recorded from the scalp with an electrode at

the vertex of the head (Cz or Fz) referenced to an electrode in the vicinity of the stimulated ear (mastoid or earlobe). To extract the ABR from the auditory evoked potential, a frequency band-pass filter from 30 to 3000 Hz is necessary, and, due to the small amplitude of the components, an average of at least 2000 stimulus presentations is required. The sampling rate of the electroencephalographic signal should not be lower than 20 kHz to avoid waveform distortions (Picton 2011; Pratt 2012). When measuring ABRs, the stimulus polarity must also be taken into account, as although the full set of components can be obtained with stimuli presented in any polarity (i.e., condensation and rarefaction), the latency of the different components is modulated depending on it (Ballachanda et al. 1992; Hall 2007).

A range of non-pathological factors affect the transient ABR, such as subject's age, gender, and body temperature. Also, stimulus factors like intensity or presentation rate do affect the recorded ABR (Picton 2011; Pratt 2012); a typical example is the shortening of the ABR peak latencies by increasing stimulus intensity. On the other hand, the ABR has high sensitivity, specificity, and reliability, and it is not susceptible to the evoking stimulus being attended or ignored nor to changes during sleep or under anesthetics (Picton 2011). Fields of clinical application of ABR recordings are, among others, newborn and infant auditory screening (i.e., the estimation of auditory sensitivity in infants and difficult-to-test children) and neurodiagnosis of eighth nerve or auditory brainstem dysfunction (i.e., monitoring eighth nerve and auditory brainstem status intraoperatively during surgery potentially affecting the auditory system and the diagnosis of auditory neuropathy) (Hall 2007).

Frequency-Following ABRs

The frequency-following response (FFR) is a component of the auditory brainstem response that reflects sustained and synchronous neural phase-locking to the individual cycles of the eliciting stimulus waveform and/or to the periodicity in the envelope of an acoustic stimulus (Skoe and Kraus 2010; Kraus et al. 2017). The scalp-recorded FFR can be evoked by frequencies in the range of 100–1500 Hz approximately. By reflecting the phase-locked activity to sounds, the FFR physically "mimics" the eliciting stimulus and is as complex as it, thus providing a stable window into the neural transcription of sounds in neuronal aggregates along the auditory hierarchy (Galbraith et al. 2000; Bidelman 2018).

The FFR was first recorded in humans by Moushegian, Rupert, and Stillman in 1973 (Moushegian et al. 1973), and it is suggested to represent phase-locked neural activity from multiple generators within the auditory system. Currently, the neural origins of FFR are still debated, yet it is treated as a putative window into subcortical sound encoding. The FFR can be obtained under passive and active listening paradigms, and it is highly sensitive to context-dependent contingencies and to real-time statistical properties of the stimulation sequence. It is also modulated by short-term auditory training and by the individual's lifelong auditory experience, such as that with language and music. Consequently, the FFR has become a relevant tool in assessing the neural encoding of speech sounds in both healthy and clinical populations (Escera 2017; Kraus et al. 2017).

By means of the appropriate analytical tools, the FFR provides an objective indicator of fundamental acoustic features intrinsic to speech and other complex sounds, including timing (onsets), pitch (fundamental frequency, f0) and timbre (the harmonics information), as well as their timing (see Kraus et al. 2017).

For more detailed information, cf. ▶ "Auditory Frequency-Following Responses."

Cross-References

▶ Auditory Event-Related Potentials
▶ Auditory Frequency-Following Responses

References

Ballachanda BB, Moushegian G, Stillman RD (1992) Adaptation of the auditory brainstem response: effects of click intensity, polarity, and position. J Am Acad Audiol 3:275–282

Bidelman GM (2018) Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. NeuroImage 175:56–69
Escera C (2017) The role of the auditory brainstem in regularity encoding and deviance detection. In: Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN (eds) The frequency-following response. Springer handbook of auditory research, vol 61. Springer, Cham, pp 101–120
Fobel O, Dau T (2004) Searching for the optimal stimulus eliciting auditory brainstem responses in humans. J Acoust Soc Am 116:2213–2222
Galbraith GC, Threadgill MR, Hemsley J, Salour K, Songdej N, Ton J, Cheung L (2000) Putative measure of peripheral and brainstem frequency-following in humans. Neurosci Lett 292:123–127
Gelfand SA (2010) Hearing: an introduction to psychological and physiological acoustics, 5th edn. Informa Healthcare, London
Hall JW (2007) New handbook of auditory evoked responses. Pearson Education, Boston
Jewett DL, Williston JS (1971) Auditory-evoked far fields averaged from the scalp of humans. Brain 94:681–696
Jewett DL, Romano MN, Williston JS (1970) Human auditory evoked potentials: possible brain stem components detected on the scalp. Science 167:1517–1518
Kraus N, Anderson S, White-Schwoch T (2017) The frequency-following response: A window into human communication. In: Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN, (eds) Springer handbook of auditory research, vol 61. Springer International Publishing, Cham
Maloff ES, Hood LJ (2014) A comparison of auditory brain stem responses elicited by click and chirp stimuli in adults with normal hearing and sensory hearing loss. Ear Hear 35:271–282
Moushegian G, Rupert AL, Stillman RD (1973) Scalp-recorded early responses in man to frequencies in the speech range. Electroencephalogr Clin Neurophysiol 35:665–667
Picton TW (2011) Human auditory evoked potentials. Plural Publishing, San Diego
Pratt H (2012) Sensory ERP components. In: Luck SJ, Kappenman ES (eds) Oxford handbook of event related potential components. Oxford University Press, Oxford, pp 89–114
Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: a tutorial. Ear Hear 31:302–324
Starr A, Zeng F, Michalewski H, Moser T (2008) Perspectives on auditory neuropathy: disorders of inner hair cell, auditory nerve, and their synapse. In: Dallos P, Oertel D, (eds) The senses: a comprehensive reference. Elsevier, Amsterdam, pp 397–412
Stone JL, Calderon-Arnulphi M, Watson KS, Patel K, Mander NS, Suss N, Fino J, Hughes JR (2009) Brainstem auditory evoked potentials – a review and modified studies in healthy subjects. J Clin Neurophysiol 26:167–175

Auditory Cortex: Separating Signal from Noise

Brian Malone
Department of Otolaryngology and Head and Neck Surgery, University of California, San Francisco, CA, USA

Synonyms

Auditory scene analysis; Auditory scene segmentation

Definition

The process of separating signal from noise is an aspect of the process of auditory scene segmentation in which the pressure waves from multiple, concurrent sound sources that mix in the air must be de-mixed by perceptual processes to generate an auditory object-based model of sound sources in the environment. Sounds that interfere with the detection or discrimination of an informative sound of interest, the signal, are often described as noise. The use of the term "noise" is informed by a long history of psychophysical experiments that measure the adverse effects of adding sounds, usually based on stochastic processes, to more behaviorally relevant signals, such as speech. Neurophysiological experiments have demonstrated that brain regions that process sound act to filter out noise while retaining information about behaviorally important sounds, such that the neural representations of sound mixtures in central structures like auditory cortex are significantly de-noised relative to representations of those sounds in more peripheral brain regions, such as the cochlear nucleus. The effects of background noise on neural responses to a signal depend on the relative amplitudes of the signal and noise, which is typically quantified in terms of the signal to noise ratio (SNR), in decibels. The similarity between signal and noise in terms of frequency content, temporal pattern, and spatial location have important implications for how

much noise at a particular SNR will disrupt processing of the signal. Neurophysiological recordings from auditory cortex have demonstrated significant diversity in how background noise affects cortical responses to certain classes of signals, including cases where noise at moderate SNRs can enhance the cortical representation of some signals.

Detailed Description

Perceptual segmentation of complex scenes relies on assigning the most likely distribution of sound sources based on the mixtures of sounds arriving at the ears (Bregman 1990; Yost 1992). Most naturally occurring sounds are complex and time-varying, so the auditory system must not only correctly attribute acoustic energy at different frequencies to a single sound source, but, in the presence of multiple concurrent sounds, must often parse acoustic energy at a single frequency among multiple sources. When thinking about this process, it is often useful to categorize sounds as "signal" or "noise" with respect to the listener's specific goals. For example, on a busy street, the sound of passing cars is noise that interferes with the conversation between two pedestrians. However, the pedestrians' conversation can also be noise that interferes with the ability a nearby bystander listening for the sound of passing cars to determine whether it is safe to cross the street. The term "background noise" is often contrasted with "foreground sound" to indicate the attentional focus of the listener.

Although the neurophysiological literature sometimes makes reference to intrinsic noise that results in variable neural responses across repeated presentations of identical stimuli, for the purposes of this entry, "noise" should be understood to refer to external, physical sounds that disrupt the processing of another sound. Traditionally, the sounds used as background noises in studies of cortical responses tend to be stochastic in nature, and as such, are defined by their time-averaged statistics, rather than a specific time-frequency pattern. For example, broadband or "white" noise has a flat spectrum over the audible range when measured over a sufficiently long interval, but the spectra of short segments of white noise vary considerably, and the output of a filter with a narrow bandwidth will show considerable modulation over time. These considerations are important when considering the expected responses of cortical neurons. In primary auditory cortex, for example, sharply tuned neurons with short temporal integration windows can exhibit highly reproducible, time-locked responses to repeated presentations of an identical ("frozen") instance of broadband noise (Scott et al. 2011). In fact, humans in laboratory settings distinguish among distinct instances of broadband noise (Agus et al. 2010). In natural listening conditions, however, the exact time-frequency patterns of environmental noise sources such as wind or ocean waves do not repeat and are not perceptually meaningful beyond their adverse effects on the processing of more behaviorally relevant signals (McDermott et al. 2013).

We define the neural representation of a stimulus as the pattern of activity that occurs in response to the stimulus and conveys information about that stimulus to auditory structures further along in the processing pathway. Background noise can corrupt the signal representation by eliciting responses that would otherwise not occur, or by reducing or eliminating responses that are typically elicited by the signal. Effectively, signals and noises compete for neural representation by influencing the patterns of neural activity in a given auditory area. It is possible that changes in signal representations might not necessarily interfere with perceptual discrimination if those changes are small relative to the differences in the responses to a set of distinct signals. Central auditory representations of complex sounds have been described as "noise invariant" (Moore et al. 2013; Rabinowitz et al. 2013) when response features associated with the signals are preserved while those associated with the noise are attenuated at successive stages of the auditory pathway.

Early studies that measured auditory cortical responses to simple stimuli such as pure tones containing a single frequency tended to focus on how tuning curves that relate neural firing rates to signal amplitude shift in the presence of additional background noise (Ehret and Schreiner 2000;

Phillips 1990; Phillips and Cynader 1985; Phillips et al. 1985) and have typically reported increases in response thresholds and latencies. Studies that have employed complex, time-varying signals have considered how the pattern of cortical responses changes in the presence of noise. For example, background noise reduces synchronization to the phrases of primate vocalizations (Nagarajan et al. 2002) and birdsong (Narayan et al. 2007).

Despite increasing research focus on cortical signal in noise processing (Bar-Yosef and Nelken 2007; Schneider and Woolley 2013), important questions remain about the underlying neural mechanisms. Proposed mechanisms include spectrotemporal filtering (Lee and Theunissen 2015), adaptation (Rabinowitz et al. 2013), synaptic depression (Mesgarani et al. 2014), and feedback gain normalization (Mesgarani et al. 2014). Recent studies have demonstrated additional complexities in how cortical neurons encode signals in noise. Similar to perceptual studies, the adverse effects of background noise at a given SNR depend on the absolute levels of both the signal and noise, such that loud noise is more disruptive than moderate noise at equivalent SNRs (Teschner et al. 2016). Among distinct clusters of cortical neurons, background noise can disrupt, fail to affect, or even enhance the representation of frequency-modulated sweeps, even in animals that are not required to attend to the stimulus (Malone et al. 2017).

Importantly, auditory cortex spans multiple levels of processing in the auditory pathway. In humans and many other vertebrate animals, the auditory cortex is comprised by a number of areas in the temporal lobe of the brain. In primates, including humans, auditory cortex is organized into core areas which receive direct projections from the ventral division of the medial geniculate body and then project to belt areas that project, in turn, to parabelt areas. Because cortical responses are believed to be more modulated by attentional mechanisms and more plastic in response to sensory exposure and perceptual learning than more peripheral structures, the question of how auditory cortical regions represent signals of interest in the presence of competing sounds is of special relevance. The balance of "bottom-up" mechanisms acting on sensory input and "top-down" mechanisms engaged by attentional focus likely shifts at different levels of the cortical processing hierarchy (Niwa et al. 2013, 2015).

Separation of signal from noise is of great clinical relevance, since the inability to follow ones stimulus among many, the "cocktail party problem" (Cherry and Bowles 1960; Cherry and Wiley 1967; Plomp 1977) is a hallmark of hearing loss, and among the most common hearing complaints in the elderly. Neurophysiological experiments in older animals suggest that reductions in the strength of inhibition may prevent the appropriate suppression of background noise (Presacco et al. 2016; Recanzone 2018).

References

Agus TR, Thorpe SJ, Pressnitzer D (2010) Rapid formation of robust auditory memories: insights from noise. Neuron 66:610–618

Bar-Yosef O, Nelken I (2007) The effects of background noise on the neural responses to natural sounds in cat primary auditory cortex. Front Comput Neurosci 1:3

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge

Cherry C, Bowles JA (1960) Contribution to a study of the "Cocktail party problem". J Acoust Soc Am 32:884

Cherry C, Wiley R (1967) Speech communication in very noisy environments. Nature 214:1164

Ehret G, Schreiner CE (2000) Regional variations of noise-induced changes in operating range in cat AI. Hear Res 141:107–116

Lee T, Theunissen F (2015) A single microphone noise reduction algorithm based on the detection and reconstruction of spectro-temporal features. Proc R Soc A 471:20150309

Malone BJ, Heiser MA, Beitel RE, Schreiner CE (2017) Background noise exerts diverse effects on the cortical encoding of foreground sounds. J Neurophysiol 118(2):1034–1054

McDermott JH, Schemitsch M, Simoncelli EP (2013) Summary statistics in auditory perception. Nat Neurosci 16:493–498

Mesgarani N, David SV, Fritz JB (2014) Mechanisms of noise robust representation of speech in primary auditory cortex. Proc Natl Acad Sci U S A 111(18):6792–6797

Moore RC, Lee T, Theunissen FE (2013) Noise-invariant neurons in the avian auditory cortex: hearing the song in noise. PLoS Comput Biol 9(3):e1002942

Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM (2002) Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. J Neurophysiol 87:1723–1737

Narayan R, Best V, Ozmeral E, McClaine E, Dent M, Shinn-Cunningham B, Sen K (2007) Cortical interference effects in the cocktail party problem. Nat Neurosci 10:1601–1607
Niwa M, Johnson JS, O'Connor KN, Sutter ML (2013) Differences between primary auditory cortex and auditory belt related to encoding and choice for AM sounds. J Neurosci 33:8378–8395
Niwa M, O'Connor KN, Engall E, Johnson JS, Sutter ML (2015) Hierarchical effects of task engagement on amplitude modulation encoding in auditory cortex. J Neurophysiol 113:307–327
Phillips DP (1990) Neural representation of sound amplitude in the auditory cortex: effects of noise masking. Behav Brain Res 37:197–214
Phillips DP, Cynader MS (1985) Some neural mechanisms in the cat's auditory cortex underlying sensitivity to combined tone and wide-spectrum noise stimuli. Hear Res 18:87–102
Phillips DP, Orman SS, Musicant AD, Wilson GF (1985) Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. Hear Res 18:73–86
Plomp R (1977) Acoustical aspects of cocktail parties. Acustica 38:186–191
Presacco A, Simon JZ, Anderson S (2016) Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. J Neurophysiol 116(5):2346–2355
Rabinowitz NC, Willmore BD, King AJ, Schnupp JW (2013) Constructing noise- invariant representations of sound in the auditory pathway. PLoS Biol 11:e1001710
Recanzone G (2018) The effects of aging on auditory cortical function. Hear Res. 2018 Sep; 366:99–105. https://doi.org/10.1016/j.heares.2018.05.013. Epub 2018 May 23
Schneider DM, Woolley SM (2013) Sparse and background-invariant coding of vocalizations in auditory scenes. Neuron 79:141–152
Scott BH, Malone BJ, Semple MN (2011) Transformation of temporal processing across auditory cortex of awake macaques. J Neurophysiol 105:712–730
Teschner MJ, Seybold BA, Malone BJ, Hüning J, Schreiner CE (2016) Effects of signal-to-noise ratio on auditory cortical frequency processing. J Neurosci 36:2743–2756
Yost WA (1992) Auditory perception and sound source determination. Curr Dir Psychol Sci 1(6):179–184

Further Reading

The Auditory System at the Cocktail Party, Springer Handbook of Auditory Research. https://www.springer.com/gp/book/9783319516608

Auditory Event-Related Potential (AERP)

► Auditory Event-Related Potentials

Auditory Event-Related Potentials

Istvan Winkler[1], Susan Denham[2] and Carles Escera[3,4,5,6]
[1]Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, MTA, Budapest, Hungary
[2]Cognition Institute and School of Psychology, Plymouth University, Plymouth, Devon, UK
[3]Brainlab-Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain
[4]Institute of Neurosciences, University of Barcelona, Barcelona, Spain
[5]Institut de Recerca Sant Joan de Déu (IRSJD), Barcelona, Spain
[6]Institute for Brain, Cognition and Behavior (IR3C), University of Barcelona, Barcelona, Spain

Synonyms

Auditory event-related potential (AERP); Auditory evoked field (AEF); Auditory evoked potential (AEP)

Definition

Auditory event-related potentials are electric potentials (AERP, AEP) and magnetic fields (AEF) generated by the synchronous activity of large neural populations in the brain, which are time-locked to some actual or expected sound event.

Detailed Description

Measurement and Derivation of AERPs/AEFs

AERPs are derived from the continuous electro−/magnetoencephalogram (EEG/MEG) by extracting segments of the signal (epochs) time-locked to some actual or expected acoustic event. AERPs were first recorded by Hallowell and Pauline

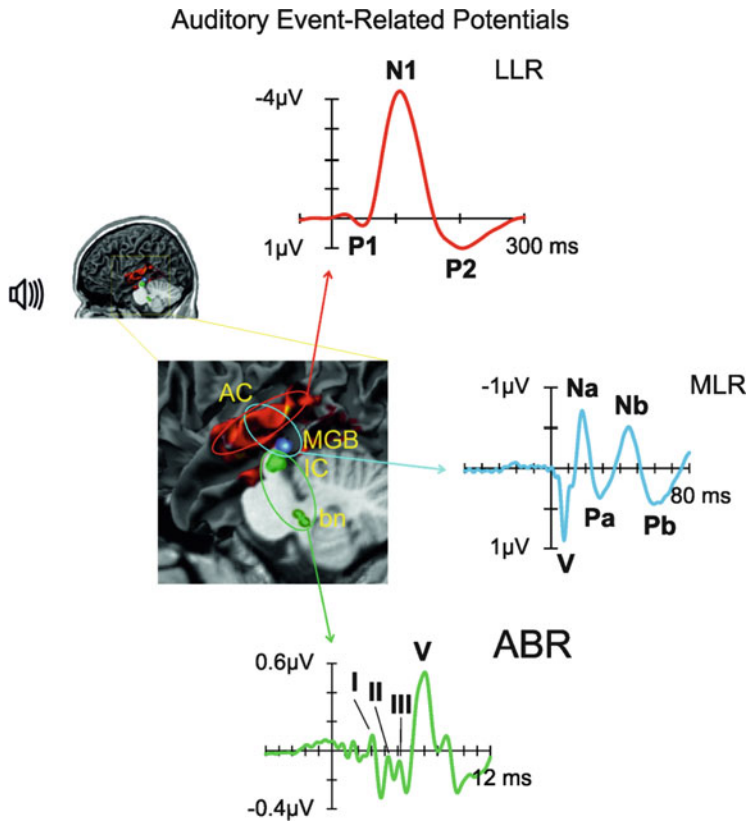
A. Davis in 1935–1936 (Davis 1939; Davis et al. 1939). Because EEG/MEG is typically recorded non-invasively (outside the brain, e.g., from/around the scalp), these measures only reflect synchronous activity of large neural populations (for measuring methods and instrumentation). Consequently, the acoustic events eliciting detectable AERPs consist of relatively large changes of spectral energy occurring within a relatively short time period, such as abrupt sound onsets, offsets, and changes within a continuous sound, because large acoustic changes affect many neurons within the auditory system and the short transition period synchronizes the responses of individual neurons (Nunez and Srinivasan 2006; cf. ▶ "Anatomy and Physiology of the Mammalian Auditory System"). Furthermore, the expectation of such changes in the auditory input can elicit AERP responses even in the absence of actual stimulation (cf. the *Omitted Stimulus Response* in "Long-Latency AERP Responses," below).

The EEG/MEG signal mixes together on-going (spontaneous) neuroelectric activity with that elicited by the event. In order to better estimate the brain activity evoked by the event, it is usually repeated several times (typically 50–200 trials/sweeps, but up to 2000 times for ▶ "Auditory Evoked Brainstem Responses") and the EEG/MEG segments are entered into some mathematical algorithm extracting the common part of the single-trial epochs. The most commonly used method for extracting AERPs aligns the single-trial epochs by their common onset and averages them point by point (the averaging method; Alain and Winkler 2012). There are many other algorithms for extracting ERPs from EEG/MEG, each based on different assumptions regarding the properties of the event-related response and the spontaneous EEG/MEG activity (for a general primer, see Luck 2005; for detailed discussion of ERPs, see Handy 2005; Fabiani et al. 2007; for special considerations of MEG/AEFs, see Hansen et al. 2010; Nagarajan et al. 2012; for AERPs, see Picton 2010; Alain and Winkler 2012).

EEG/MEG signals can contain components up to a few kHz with the faster components mainly originating from lower levels of the auditory system (cf. ▶ "Anatomy and Physiology of the Mammalian Auditory System" and ▶ "Auditory Evoked Brainstem Responses"). Cortical contributions are much slower, up to a few tens of Hz. Unless one is specifically looking for very slow (Vanhatalo et al. 2005) or fast responses (Curio 2005), AERP recordings are usually made with bandpass filter settings of 0.01–50 Hz (or 250 Hz for extracting the Middle-Latency Response, see below). AERP amplitudes are typically below 10 μV with the reference (zero) level set to a baseline voltage (unless direct current is recorded), which is usually the average signal amplitude in a time interval preceding the AERP-eliciting event. Although in general, there is no unique solution to the inverse problem of finding the origins (the neural generators) of electromagnetic potentials measured outside the brain, by utilizing anatomy/physiology based constraints, the generators of AERPs can be located with reasonable accuracy in the brain (Nunez and Srinivasan 2006; see also ▶ "Brain Imaging: Overview"). Due to the underlying physics, magnetic AEFs provide more accurate source localization compared with electric AERPs (Nunez and Srinivasan 2006; Nagarajan et al. 2012). On the other hand, MEG only allows one to measure the tangential components of the electromagnetic activity in the brain, whereas EEG represents the full activity (Hansen et al. 2010). For AEFs, however, this limitation of the MEG signal is less severe than for other sensory/cognitive systems (Picton 2010; Nagarajan et al. 2012). This is because a large part of the human auditory system in the cortex is located in the Sylvian fissure (see ▶ "Anatomy and Physiology of the Mammalian Auditory System"), thus mostly producing magnetic signals which can be picked up by the MEG device.

AERP Waves, Components, Naming Conventions

Figure 1 illustrates the progression of stimulus-related neuronal activity through the auditory system and the corresponding series of positive and negative waveforms observable in the AERP response. The earliest detectable responses (< ca. 10 ms after the acoustic event) originate



Auditory Event-Related Potentials, Fig. 1 The human Auditory Event-Related Potential (*AERP*), its main waveforms and its generators in the brain. The human AERP is composed of three groups of waveforms in three different latency ranges: the Auditory Brainstem Response (*ABR*) elicited within the first 8–10 ms from sound onset (*green*, *bottom panel*); the Middle-Latency Response (*MLR*), elicited within the 12–50 ms interval from sound onset (*blue*, *central panel*), and the Long-Latency Responses (*LLR*) emerging after 50 ms (*red*, *top panel*). The anatomical inset (*left panel*) highlights the main stages of the auditory pathway: "*bn*", brainstem nuclei (including the cochlear nucleus, the superior olivary nucleus, the nucleus of the lateral lemniscus); "*IC*", inferior colliculus; "*MGB*", the medial geniculate body in the thalamus; "*AC*", auditory cortex. The main assumed brain sources of the different AERPs are marked by colored circles: the ascending auditory pathway of the brainstem for ABRs (*green*); the thalamo-cortical loops and parts of auditory cortex for MLRs (*blue*); the auditory cortex for LLRs (*red*). AERPs can be broken down into a series of waves (see the naming convention in the main text)

from subcortical brain structures and are termed the *Auditory Brainstem Response* (ABR; cf. ▶ "Auditory Evoked Brainstem Responses"). These are followed by AERP responses of thalamo-cortical origin (mainly from the primary auditory cortex), termed the *Middle-Latency Response* (MLR), elicited during the ca. 10–50 ms post-event latency range. The waveforms following are called *Long-Latency Responses* (LLR) and they originate largely from auditory cortex, but may also include contributions from parietal and frontal areas.

ABRs are referred to by Roman numerals set in the order of their elicitation. MLR waveforms are usually denoted by their polarity at the vertex (approximately the top of the head); P for positive and N for negative polarity waves, and a letter or a number (see Fig. 1). There are two conventions for the numbers in referring to LLRs: They either denote the serial order of the response starting with the first detected response (Davis 1939; Davis et al. 1939), termed N1, or they denote the typical peak latency of the waveform, such as P50 (the same as Pb or P1). However, as more and

more responses elicited with the same polarity and in overlapping latency ranges have been discovered, both notations have become equivocal. Therefore, some recently discovered AERP responses are denoted by acronyms referring to their functional aspects, such as ORN (Object Related Negativity) or MMN (Mismatch Negativity) (for a detailed description of the variety of ERP responses, see Luck and Kappenman 2012; for AERPs, Picton 2010; Alain and Winkler 2012). Magnetic response fields are usually marked by the letter 'm' appended to the name of the corresponding (A)ERP (e.g., N1m or N100m).

Beyond the categorization based on the ERP peak latency there are two other typical distinctions in use. ERPs are termed obligatory or exogenous if they are elicited by each event irrespective of its relation to preceding or concurrent events or the person's task, motivations, knowledge, etc. ERP components elicited only when there is a certain relation between the event and other events or some aspect of the person's mental state are termed endogenous. Another distinction refers to the person's voluntary activity with respect to the given stimulus event. ERP responses only elicited when the person has some explicit task involving the event (task-relevant even) are termed "active" ERP responses, while those elicited irrespective of the person's task (task-irrelevant) are termed "passive" ERP responses.

However, waveforms (peaks and dips) are not the true building blocks of ERP responses. The brain is a massively parallel processing instrument. Therefore, at any given moment of time, multiple processes may contribute to the observable waveform. For a neurophysiologically and functionally more meaningful decomposition of the complex neuroelectric response, one should be able to delineate how each of the concurrent processes contributed to the observed neuroelectric activity. This objective is reflected by Näätänen and Picton's (1987) definition of an ERP component: '. . . we define an EP "component" as the contribution to the recorded waveform of a particular generator process, such as the activation of a localized area of cerebral cortex by a specific pattern of input' (p. 376). Thus a component is defined by two criteria: (1) it should have a specific generator structure (e.g., secondary auditory and frontal cortices) and (2) it should be specific to some experimentally definable stimulus configuration (such as stimulus change after several stimulus repetitions). One could amend this definition with the person's task/goals/knowledge regarding the given stimulus configuration (e.g., instructed to respond to the given stimulus event). However, the criteria set up by the above definition are seldom met in ERP research. This is partly due to limitations in separating generators (i.e., they are usually distributed over an area in the brain and concurrently active processes often occupy areas very close, possibly even overlapping each other) as well as not knowing what stimulus configurations are handled by the same processes in the brain (are all expectation violations processed in the same way? – probably not). Thus in practice, the majority of ERP research reports use the terms "waveform" and "component" interchangeably, sometimes linking the effects of multiple manipulations to the same waveform, while at other times, attempting to separate the specific generator process affected by a given stimulus or state variable.

There are many different processes, which can be reflected in AERPs. Early, obligatory responses typically reflect processes extracting auditory features, such as pitch, intensity, location, etc. Most AERP responses are sensitive to the amount of sound energy change and also to some aspects of the sound presentation rate or the ratio between sound and silence in time. These attributes of auditory stimuli belong to the primary descriptors of sound events as studied in psychoacoustics (Zwicker and Fastl 1990). There are also AERP responses indicating the presence of automatic memory for sounds (Cowan 1984; Demany and Semal 2007) and predictive processing of the auditory input (Friston and Kiebel 2009; Winkler et al. 2009). Further, some AERP responses reflect processes involved in auditory scene analysis (Bregman 1990), the separation of concurrently active sound sources in the environment and the formation of auditory perceptual objects (Griffiths and Warren 2004; Winkler

et al. 2009). Many AERP responses are also sensitive to attentional manipulations, including the active storage of sounds, selective attention, and target identification (Cowan 1988; Näätänen 1990). AERP responses specific to music and speech perception are described in the corresponding entries (► "Music Processing in the Brain" and ► "Electrophysiological Indices of Speech Processing"). Therefore, AERPs have been extensively used to test theories of perception (e.g., Bregman 1990; Friston 2005), memory (e.g., Broadbent 1958; Baddeley and Hitch 1974; Cowan 2001), and attention (e.g., Broadbent 1958; Lavie 1995) and in recent years they have received increased interest from computational modelling (e.g., Garrido et al. 2009; May and Tiitinen 2010; Wacongne et al. 2011) as well as from clinical applications (e.g., Picton 2010; Näätänen et al. 2012).

In the following, we shall describe the most important middle- and long-latency AERP responses (for the auditory brainstem responses, see ► "Auditory Evoked Brainstem Responses").

Middle-Latency AERP Responses

Discrete auditory stimuli elicit a sequence of very small (<1 μV) negative and positive waveforms in the 10–50 ms post-stimulus latency range, termed the Middle Latency Response (MLR). These responses can usually be best seen on signals recorded from the vertex with a mastoid or neck electrode as reference. The names and typical latencies of MLRs when elicited by click stimuli are: *N0* (10 ms), *P0* (15 ms), *Na* (20 ms), *Pa* (30 ms), and *Nb* (~40 ms) (see Picton 2010). An additional later waveform, the *Pb*, which peaks at about 50 ms from sound onset, is not always included amongst the MLR components, because it can also be obtained as the P50 or P1 with the filter bandwidth optimised for measuring LLRs (Regan 1989; see below). Because of their small amplitude and specific spectro-temporal characteristics, recording the MLR requires (a) averaging across close to 1000 responses, (b) appropriate filter settings (15–200 Hz, Bell et al. 2004), and (c) careful removal of electromagnetic interference from power supplies and lines, as a large part of the power of the MLR responses falls into the 50–60 Hz range. It is also important to avoid artefacts stemming from the myoelectric activity of the postauricular muscle (PAM), which lies behind the ear and is activated by loud sounds. This is usually achieved by placing the reference electrode on the neck or the sternum (Bell et al. 2004). Optimal sounds for eliciting clear MLRs are chirps and clicks, which have sharp onsets and a broad spectrum. Pure tones elicit MLRs of somewhat different morphology and smaller amplitude (Borgmann et al. 2001). However, MLRs can be obtained even with low-intensity tone bursts and relatively independently of the arousal level (Jones and Baxter 1988).

No hemispheric asymmetry was found for MLRs as a function of the stimulated ear (Starr and Don 1988). Based on precise structural maps of individual brains, the spatiotemporal pattern of neural activation giving rise to MLRs has been identified in supratemporal auditory areas using either current estimates derived from intracerebral recordings (Yvert et al. 2005) or equivalent dipole source modelling of scalp-recorded electric brain potentials (Yvert et al. 2001). These studies localized the earliest cortical activity (P0) at 16–19 ms from sound onset in the medial portions of Heschl's sulcus (HS) and Heschl's gyrus (HG), which likely correspond to primary auditory cortex (PAC). Na generation resulted from activity in more posterior regions of the same HS and HG areas. During the Pa/Pb complex, which includes also the Nb, the electric brain activity propagates in postero-anterior and medio-lateral directions in HG to the Planum Temporale (PT) and then to more anterior parts of the Superior Temporal Gyrus (STG), which correspond to secondary auditory areas. Also, frontal and parietal brain regions contribute as early as 30 ms from sound onset (the P30 m AEF response) to MLR (Itoh et al. 2000). Animal studies have suggested that MLRs involve parallel thalamocortical activation of areas 41 (PAC), and 36 (parahippocampal gyrus), while human lesion studies have implicated contributions from thalamic projections to Pa (Kraus et al. 1982) and Na (Kaseda et al. 1991),

supporting a thalamo-cortical interaction in MLR generation.

With increasing sound intensity, MLR component latencies decrease while the amplitudes increase, although these effects may not uniformly apply to each component (e.g., Na, but not Pa; Seki et al. 1993; Althen et al. 2011). Galambos et al. (1981) found a systematic reversed U-shaped relationship between the MLR amplitudes and stimulus presentation rate. At slow rates (≤10 Hz), peak-to-trough amplitudes are rather small (0.4 μV) and they reach the maximum of 1 μV by about 40 Hz presentation rate. This twofold increase in amplitude is due to superimposition of MLRs elicited by successive sounds. In contrast, at stimulation rates below and above 40 Hz out-of-phase responses to successive MLR responses cancel out each other. Some authors interpret this finding in terms of the "steady state" potentials (oscillatory activity generated in sensory cortical areas that is time-locked to the periodicity of stimulus presentation; typically measured from visual and somatosensory cortical areas; Rees et al. 1986). Other authors assume that this phenomenon reflects the contribution of transient early evoked gamma-band oscillations to the auditory MLR (Basar et al. 1987; Pantev et al. 1991; Müller et al. 2001). Based on the stimulus-driven properties outlined above, MLRs have been considered exogenous AERP components. However, this view has been challenged by studies showing that MLRs are enhanced by strongly focused attention as early as 20 ms from sound onset (Woldorff and Hillyard 1991; Woldorff et al. 1993; cf. "Attention-Related AERP Responses" below), and that MLR amplitudes are modulated as early as 50 ms from sound onset by task difficulty and whether or not a motor response is required (Ninomiya et al. 1997). Further, a recent series of studies has shown that MLRs are sensitive to stimulus probability in a feature-specific manner (Grimm and Escera 2012) with infrequent frequency changes enhancing the Pa (Slabu et al. 2010) and Nb (Grimm et al. 2012; Alho et al. 2012), whereas location changes enhance the Na (Sonnadara et al. 2006; Grimm et al. 2012; Cornella et al. 2012). These results suggest that the MLR components reflect processes subserving higher-order sensory/cognitive functions.

Long-Latency AERP Responses

The auditory *P1* (*P50, Pb*; Fig. 1) component is at the border between MLR and LLR. In fact, when recorded and analysed with the filter setting most useful for deriving MLRs it is termed the Pb (see "Middle-Latency AERP Responses," above). Using the parameters better suited for assessing LLRs, it typically peaks at about 50 ms from stimulus onset, appearing with positive polarity at the vertex and with reversed (negative) polarity at electrodes placed on the other side of the Sylvian fissure (e.g., electrodes placed over the mastoid apophysis). P1 is the first wave of the P1-N1-P2 obligatory exogenous AERP complex. It is thought to be generated bilaterally in primary auditory cortex, somewhat larger contra- than ipsilaterally for pure tones (Godey et al. 2001) and for other types of pitch-evoking sounds (Butler and Trainor 2012), with some spreading of the neuroelectric activity over its time course (Yvert et al. 2005). P1 is often used as a landmark for primary auditory cortex in AERP and AEF studies aimed at localizing the AERP components. Similarly to other obligatory AERP responses, P1 is highly sensitive to stimulus features and presentation rate (fully recovering within a few hundred milliseconds) as well as to attentional manipulations (Picton 2010). The P1 was initially assumed to reflect neural activity involved in extracting auditory features (e.g. Näätänen and Winkler 1999). Recent evidence also links this response with the automatic separation of auditory streams (Gutschalk et al. 2005; Snyder et al. 2006; Szalárdy et al. 2013; cf. ▶ "Auditory Perceptual Organization"): The amplitude of the P1 component has been found to be modulated by whether a sequence with two interleaved sounds (e.g., ABABAB. . ., where 'A' and 'B' denote two different sounds) was perceived as a single coherent stream or in terms of two concurrent streams of sound (one made up of the 'A' and the other by the 'B' sounds).

The auditory *N1* (*N100*; Fig. 1) wave was the first AERP response discovered historically (Davis et al. 1939) as it is the most prominent

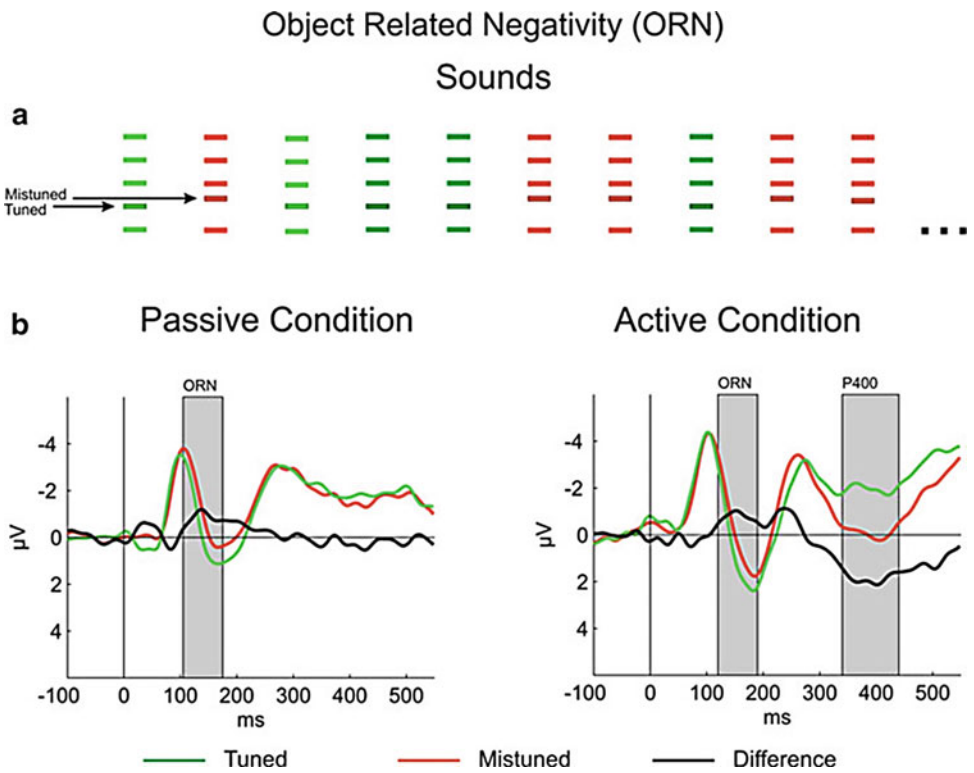
deflection at the vertex. It is elicited by abrupt changes in sound energy, such as sound onsets and offsets (Näätänen and Picton 1987). N1 typically peaks with negative polarity over the vertex ca. 100 ms after the eliciting event. It is also the most widely studied AERP response, having been linked with virtually any and all assumed auditory processing steps. The N1 wave has a complex generator (and thus subcomponent) structure (Näätänen and Picton 1987). The subcomponent most tightly related to auditory processes (the supratemporal N1) is mostly located in secondary auditory areas (Godey et al. 2001), but it also overlaps the areas active during the P1 component (Yvert et al. 2005). Similarly to the P1, N1 is larger contralaterally to the ear of stimulation and it is highly sensitive to stimulus features, presentation rate, and attentional manipulations. However, unlike the P1, the N1 recovery is much slower, extending beyond 10 s (Cowan et al. 1993). Further, N1 is sensitive to perceived sound features (e.g., pitch), as opposed to raw spectral parameters (such as the harmonic frequencies of a complex tone; Pantev et al. 1989b), although feature extraction is not yet complete at the time the N1 wave is elicited (Winkler et al. 1997). The supratemporal N1 also shows both tonotopic (Pantev et al. 1988) and ampliotopic organization (Pantev et al. 1989a); that is, the location of its generator varies with the frequency and amplitude of pure tones. However, the N1 generators are not sensitive to combinations of sound features (i.e., feature conjunctions).

The processes reflected by N1 have been linked with onset and acoustic change detection (Näätänen 1992), feature extraction, sensory memory (Lü et al. 1992; at least for sound features, Näätänen and Winkler 1999) and, recently, with auditory stream segregation (Gutschalk et al. 2005; Snyder et al. 2006; Szalárdy et al. 2013). For example, the length of the silent period after which an N1 with maximal amplitude is elicited by a sound is in good correspondence with the behaviourally measurable duration of auditory sensory memory traces (Cowan 1984). When sounds are presented in a train with <10 s silent intervals between them, the N1 amplitude decreases sharply within the first few presentations, reaching an asymptote within 5–10 presentations (e.g., Cowan et al. 1993). Based on this finding, some authors argue that through adaptation (see ▶ "Adaptation in Sensory Cortices, Models of"), the neurons underlying the N1 response may retain all sound information and thus provide the basis for detecting violations of auditory regularities (May and Tiitinen 2010). However, this hypothesis is debated in the literature (e.g., Näätänen et al. 2011). The sensitivity of the auditory N1 wave to selective attention initially suggested that the difference between the N1 responses elicited by task-relevant (attended) and task-irrelevant (unattended) sounds (the Nd; Hillyard et al. 1973) may reflect an orientation to the attended auditory features and/or maintenance of the memory trace of the target sound. However, others argued that the differential response is separate from the N1, with the early part overlapping the N1 (termed Nd_e) assumed to reflect feature processing, and the later part (Nd_l, also termed the Processing Negativity, PN; Näätänen 1982; see PN in "Attention-Related AERP Responses") the maintenance of the attentional trace (Koch et al. 2005; Näätänen et al. 2011).

Little is known about the auditory *P2* (*P175*, *P200*; Fig. 1) AERP response. It has been mostly studied within the P1-N1-P2 or N1-P2 complex. P2 typically peaks between 175 and 200 ms from the event onset with positive polarity over the vertex, inverting polarity over the Sylvian fissure. The generators of P2 lie anterior to those of the N1 in secondary auditory areas (Mäkelä et al. 1988; Bosnyak et al. 2004). Lesion (Woods et al. 1993) and maturation studies (Ponton et al. 2000) suggest that P2 may reflect the output of the mesencephalic reticular activating system (see ▶ "Anatomy and Physiology of the Mammalian Auditory System"). Only a few studies have attempted to distinguish P2 from the N1 wave. The P2 amplitude was found to be more sensitive to perturbing the feedback of one's voice than the N1 (Behroozmand et al. 2009) as well as to training with specific types of sounds (e.g., speech: Tremblay et al. 2001; music: Bosnyak et al. 2004; or frequency discrimination: Tong et al. 2009). There are several speculations regarding

the functions of the processes reflected by P2. Based on its assumed neural origin, P2 has been suggested to be generated by a pre-attentive alerting mechanism (Tremblay and Kraus 2002). Other suggestions include P2 reflecting stimulus classification (Crowley and Colrain 2004), modulating the threshold for conscious perception (Melara et al. 2002), protecting against interference from irrelevant stimuli (Garcia-Larrea et al. 1992), and the accuracy of memory traces in short-term memory (Atienza et al. 2002).

The *Object Related Negativity* (*ORN*) is elicited when more than one sound are simultaneously heard (Alain et al. 2001). Thus ORN reflects the outcome of the analysis of simultaneous (concurrent or vertical) auditory grouping cues (cf. ▶ "Auditory Perceptual Organization"). Components of sounds emitted by a single source usually commence at the same time, they originate from the same spatial location and, if composed of discrete frequencies, they consist of harmonics derived from the same base (i.e., integer multiples of the same frequency). When the acoustic input does not meet these criteria, one usually experiences it as two or more concurrent sounds and ORN is elicited. ORN is typically recorded by presenting complex tones with one harmonic mistuned by 4% or more (Fig. 2, panel A) and derived by subtracting the response to the one-sound stimulus (e.g., tuned tone) from that to the multiple-sound stimulus (e.g., mistuned tone). ORN peaks between 140 and 180 ms from sound onset, with



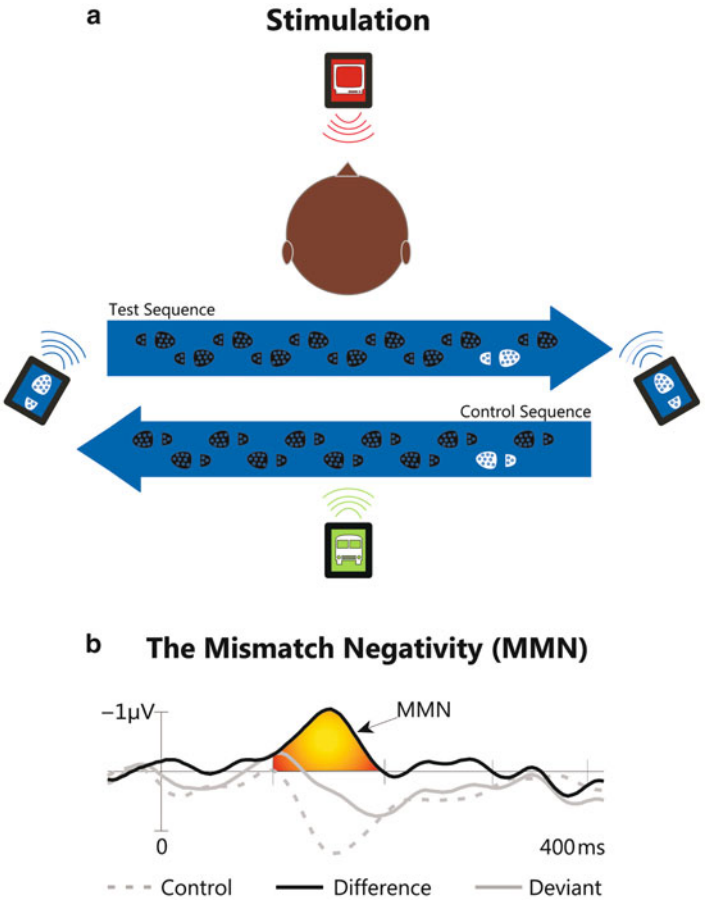
Auditory Event-Related Potentials, Fig. 2 Object Related Negativity (*ORN*) (**a**) Complex tones with the second of five harmonics tuned (*green*) or mistuned upwards by 8% (*red*) were presented equiprobably in a sequence. (**b**) Group-averaged ($N = 20$, *left*; $N = 23$, *right*) AERP responses elicited by tuned and mistuned complex tones recorded at the vertex, separately in the passive (participants disregarded the sounds) and the active condition (participants judged whether they heard one or two concurrent tones). Mistuned-minus-tuned difference waveforms (*black*) show a negative waveform appearing between 100 and 200 ms from sound onset in both task conditions. This is the ORN response (the range is marked by *grey shading*). The positive difference waveform observed in the 300–500 ms latency range in the Active Condition is termed the P400

the largest amplitude over the fronto-central region of the scalp (Fig. 2, panel B *left*). ORN has bilateral neural generators in auditory cortex, which are separate from those of the previously described obligatory AERP responses (Arnott et al. 2001). Some studies have indicated the existence of two independent lateralized generator processes, since although ORN is elicited even when most tones in the sequence have been mistuned, the probability of mistuned sounds within the sequence differentially affected the ORN generators in the two hemispheres (Bendixen et al. 2010). If the listener is instructed to respond when he/she hears two concurrent sounds, a late positive response (P400) is elicited in addition to the ORN (Fig. 2, panel B *right*; Alain et al. 2001).

The auditory *N2* (*N200*; Fig. 1) wave covers at least three (*N2a or MMN*, *N2b*, *N2c*; see Pritchard et al. 1991), possibly more AERP components (Folstein and Van Petten 2008) appearing partly overlapping in time between 150 and 300 ms from the eliciting event. The somewhat earlier N2a or MMN does not require attention to be focused on the event (cf. MMN), whereas the later components are related to attentive monitoring of the acoustic input and they are not specific to sounds.

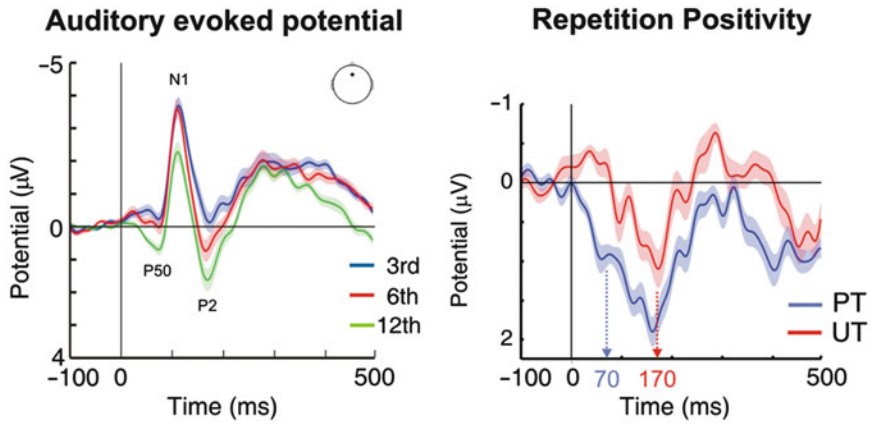
The *Mismatch Negativity* (*MMN*, *N2a*) is an AERP component elicited by violations of auditory regularities (Winkler 2007; Näätänen et al. 2011; Fig. 3). MMN typically emerges between 100 and 200 ms from the onset of deviation with frontocentrally dominant negative polarity that is inverted over the Sylvian fissure. MMN generators are located bilaterally in secondary-auditory and frontal areas (Alho 1995). Although traditionally regarded as a component reflecting auditory change detection, technically, MMN does not reflect acoustic change, as for example, an alternating sequence of sounds does not elicit the MMN, whereas repeating a sound within such a sequence does (Horváth et al. 2001). MMN is derived by subtracting from the response elicited by the regularity-violating sound (termed "deviant") the response elicited by a control sound. Optimally, the control sound is either identical or very similar to the deviant sound but does not violate any auditory regularity (for a detailed discussion of selecting the correct control, see Kujala et al. 2007). MMN is elicited even when the sounds are task-irrelevant, although it can be suppressed by strongly focusing attention on a parallel auditory channel and/or by contextual information (Sussman 2007). Initially discovered within the oddball paradigm (Näätänen et al. 1978), MMN has since been observed for violations of a large variety of abstract and complex regularities (Näätänen et al. 2001). In parallel, its interpretation shifted from MMN being an AERP correlate of auditory sensory memory (Näätänen and Winkler 1999; Cowan 1984) tasked with detecting potentially relevant events in the auditory environment (Näätänen 1992) towards the compatible but more general notion of representing a process that updates the detected auditory regularities when their predictions are not met by the incoming sound (Winkler 2007). The latter interpretation links MMN with predictive coding theories (Friston 2005; Winkler and Czigler 2012) and posits that it plays a role in auditory stream segregation (cf. ► "Auditory Perceptual Organization") by maintaining the predictive models underlying auditory perceptual objects (Winkler et al. 2009).

The *Repetition Positivity (RP)* appears as a fronto-central amplitude modulation of the P50, N1 and P2 AERP responses (Fig. 4); all three of them overlap the slow positive RP waveform so that the P50 and P2 become more positive and the N1 less negative with increasing number of repetitions of the eliciting sound (Haenschel et al. 2005; Costa-Faidella et al. 2011a, b). Similar stimulus repetition effects have been observed even at shorter latencies, during the MLR latency range (Dyson et al. 2005). The RP was first observed by Baldeweg et al. (2004) and characterized by Haenschel et al. (2005) in a study that aimed at investigating the neural correlates of the sensory memory trace implicated in the generation of the MMN. It was argued that the MMN amplitude dependence on the number of standard-stimulus repetitions preceding the deviant (e.g., Sams et al. 1983; Javitt et al. 1998) provides only an indirect measure of the strength of the underlying memory trace. The AERP elicited by the standard sound was expected to show effects of repetition suppression (Desimone 1996), as was



Auditory Event-Related Potentials, Fig. 3 The Mismatch Negativity (*MMN*). (**a**) The experimental setup. Participants watched and listened to a movie presented on a TV screen directly in front of them. A series of footsteps perceived as moving from left to right (Test Sequence; *upper arrow*) or right to left (Control Sequence; *lower arrow*) were delivered by a pair of loudspeakers placed symmetrically on two sides, slightly behind the participant's head. Ten out of the 11 different digitized natural footstep sounds (marked as *black* footprints on the *blue arrows*) could be perceived as a coherent sequence produced by someone walking across a room. The tenth footstep of the Test Sequence ("deviant") and the second footstep of the Control Sequence ("control") however sounded as if the person stepped on a different surface (marked by the *white footprint* on the *blue arrows*). Street noise was delivered through a loudspeaker placed directly behind the participant. (**b**) Group-averaged ($N = 8$) AERP responses elicited by the deviant (*continuous grey line*) and the identical control sound (*dashed grey line*) measured from the frontal midline electrode. The MMN component, derived by subtracting the control response from that to the deviant (difference: *black line*) is marked with *yellow-orange* fill in the *MMN* latency range. The results illustrate that (1) *MMN* is only elicited when a sound violates a detected regularity, as the regular progression of footsteps needed to be detected and represented by the brain before it could be violated (which could not happen if only one "regular" footstep sound preceded the different one); (2) regularities can be extracted from acoustic variance as all regular footstep sounds were acoustically different; (3) regularities are separately maintained for concurrent auditory streams, as *MMN* was elicited for deviation in the footstep stream despite the presence of two other active sound sources; and (4) *MMN* elicitation does not require one to attend the stream in which a regularity has been violated, as participants in this experiment attended the movie, not the footsteps. (Adapted from Winkler et al. 2003)

observed for individual neurons in the primary auditory cortex of the cat (Ulanovsky et al. 2003), and this could provide a more direct measure of the strength of standard-stimulus memory trace. The typical paradigm use for obtaining the RP is called the "roving-standard"



Auditory Event-Related Potentials, Fig. 4 The Repetition Positivity (*RP*). *Left:* Group-averaged ($N = 14$) frontal midline (marked on the schematic head drawing at the *top right corner*) AERPs elicited by pure tones in a roving-standard stimulus paradigm (see in the text). The panel shows AERPs (averaged across different frequencies) elicited for the 3rd (*blue*), 6th (*red*) and 12th (*green*) repetition of the same tone. Note that the positivity covering the latency range of the P50-N1-P2 waveform complex emerges at the sixth repetition and becomes more pronounced by the 12th repetition. *Right:* Difference waveforms resulting from subtracting the response to the third repetition from that to the 12th repetition under two conditions: Predictable Timing (*PT*: isochronous presentation, *blue*) and Unpredictable Timing (*UT*: the within-train inter-onset interval was varied, *red*). Note that the onset of RP is earlier (ca. 70 ms post-stimulus) for the predictable than for the unpredictable timing condition (ca. 170 ms). (Adapted from Costa-Faidella et al. 2011a)

paradigm (introduced by Cowan et al. 1993), as the classical oddball paradigm yields less clear results (Cooper et al. 2013). In the roving-standard paradigm, short trains of a repeating sound are delivered without a break with each train delivering a different sound (e.g., pure tones with different frequencies). The number of sound repetitions can also vary from train to train. To separate the RP from other concurrent AERP components, the average response elicited by the second or the third sound of the train is subtracted from that elicited by the last tone of the train. The response to the first sound of the train is not used in the subtraction, because, due to the sound change between the trains, it should elicit the MMN (Haenschel et al. 2005; Costa-Faidella et al. 2011a, b). The generator structure of the RP has not yet been fully characterized, but its early onset latency (commencing during the P50) and its long duration (ending during the P2) suggest that it may involve a distributed cortical network spanning from PAC up to auditory association areas (Baldeweg 2007). The RP has been shown to simultaneously encode repetitions over multiple time scales (Costa-Faidella et al. 2011b; Cooper et al. 2013) similarly to single neurons observed in the cat's PAC (Ulanovsky et al. 2004). In addition to stimulus repetition, the RP is also sensitive to temporal regularities, such as whether the sounds are presented isochronously or with random timing: Costa-Faidella et al. (2011a) found earlier and larger RP's for isochronous as compared with randomly timed tones in the trains. The latter result supports the predictive coding view of auditory deviance detection (Winkler 2007; Winkler and Czigler 2012), according to which detection of a regularity helps to encode the sensory memory trace of upcoming stimuli. Thus higher levels in the auditory processing hierarchy feed back to lower processing levels (Baldeweg 2006).

Auditory brain responses can also be elicited without hearing sounds. By omitting sounds from an isochronous sequence, one can record potentials time-locked to the moment when the sequence would have continued in a regular manner. The responses are termed the *Omitted Stimulus Response* (*OSR*). Some of them are elicited

even when listeners don't focus on the sounds, thus demonstrating a basic tendency of the auditory system to generate predictions for incoming sounds (Friston 2005; Winkler et al. 2009). It has been shown that when all features of the upcoming sound can be predicted from the preceding sound sequence, the OSR elicited by sound omission during the first 50 ms does not differ from the AERP elicited by the sound itself; however, when only the timing of the sound can be predicted, but not its features, the OSR starts to differ from the corresponding AERP at an earlier time (Bendixen et al. 2009). When sounds are predictably caused by some action of the listener, occasionally omitting one elicits an AERP that is initially (up to ca. 100 ms) morphologically similar to that elicited by the corresponding self-initiated sound; although the brain generators underlying the two responses partly differ from each other (SanMiguel et al. 2013). There is also an MMN-like OSR (Yabe et al. 1999). Elicitation of these responses is limited to inter-onset-intervals (IOI) shorter than ca. 200 ms (Horváth et al. 2007), except when the omitted sound is part of a pattern (Salisbury 2012). With longer IOIs, an early posterior negative (180–280 ms) response and a later anterior positive wave have been obtained (Busse and Woldorff 2003). Further, ERP responses can also be elicited by mental imagery of sounds, although the results vary somewhat with the procedure employed (Meyer et al. 2007; Cebrian and Janata 2010; Wu et al. 2011).

Attention-Related AERP Responses

Attention-related AERPs include two distinct groups of responses: those related to involuntary (passive or exogenous) attention, and those related to voluntary, mainly selective attention. Regarding involuntary attention, at least three components have to be considered. The MMN (described above), or at least its frontal component (Giard et al. 1990; Deouell et al. 1998; Escera et al. 2000a; Deouell 2007), has been associated with involuntary attention (Näätänen and Michie 1979; Näätänen 1990, 1992). Some studies have also related the activation of the supratemporal MMN generator with behavioural correlates of involuntary attention, i.e., delayed response times to target stimuli on a primary task (Yago et al. 2001). Näätänen and Michie (1979) proposed that the process generating MMN may issue a call for focal attention (Öhman 1979) upon the detection of an unexpected change in the acoustic environment. Initial supportive evidence was provided by Schröger (1996; Schröger and Wolff 1998a) and Escera et al. (1998), who introduced new auditory-auditory and auditory-visual distraction paradigms (for a more recent design, see Horváth and Winkler 2010). In these paradigms, participants are instructed to perform a primary auditory or visual task while ignoring rare task-irrelevant violations of an auditory regularity. Several studies have shown that these rare deviations prolong the reaction time and reduce the hit rate to target stimuli in the primary task (Escera and Corral 2007), thus demonstrating involuntary attention switching to the task-irrelevant deviations.

Following the MMN, AERPs recorded in the distraction paradigm display a fronto-central positive deflection ca. 250–350 ms from stimulus onset, termed the *P3a* or *novelty-P3*. P3a was first described by Squires et al. (1975) as an earlier and more frontal positive deflection compared to the later and more posterior P3b component (for a review on P3b, see Donchin and Coles 1988). Whereas P3a is elicited by rare task-irrelevant sounds, P3b is elicited by target sounds (for a detailed comparison between the P3a and P3b, see Polich 2007). P3a is also elicited by widely different and "novel" (unique, categorically different from the context) sounds (Knight 1984), hence it is sometimes referred to as the novelty-P3 (for a discussion of whether the P3a and the novelty-P3 can be considered as the same ERP component, see Simons et al. 2001). Compelling evidence linking the novelty-P3 to the orienting reflex (OR; Sokolov 1963) was obtained by Knight (1996), who found strong correlation between the novelty-P3 and one of the well-known autonomic components of OR, the galvanic skin response (GSR). The P3a is composed of two subcomponents distinctly differing in latency (early and late), scalp distribution, and

sensitivity to attentional manipulations (Escera et al. 2000a; Yago et al. 2003). Source modelling of the magnetic counterpart of P3a (P3am) elicited by auditory deviants and novel sounds has revealed a genuine auditory cortical contribution to the early part of P3a (Alho et al. 1998). Whereas the early part of the novelty-P3 appears to be insensitive to attentional manipulations (Escera et al. 1998), the later part is modulated by working memory (SanMiguel et al. 2008) and emotional load (Domínguez-Borràs et al. 2008). The early P3a is sensitive to stimulus-specific information predicting task-irrelevant auditory deviance, whereas the late P3a appears to be more closely correlated with distraction (Horváth et al. 2011). P3a is widely regarded as a correlate of attention switching (Escera et al. 2000a; Friedman et al. 2001). However, some recent studies suggested that although P3a is probably an antecedent of attention switching it can be elicited without a corresponding shift in the focus of attention (Rinne et al. 2006; Horváth et al. 2008b; Horváth and Winkler 2010; Hölig and Berti 2010).

The third involuntary attention related AERP component is the so-called *Reorienting Negativity* (*RON*), first described by Schröger and Wolff (1998b). RON is observed as a negative deflection following the P3a (Escera and Corral 2007). RON has been suggested to reflect processes of reorientation (restoring the task set of the primary task) after a distracting stimulus. RON is composed of two subcomponents (Escera et al. 2001; Munka and Berti 2006; Berti 2008) the functional characterization of which are still debated (Escera et al. 2001; Berti 2008). The cortical generators of RON are not well known. Horváth et al. (2008a) found contributions from primary motor areas to RON, suggesting that action-selection related activity plays a role in the reorientation process. Both P3a and RON as well as behavioural correlates of distraction (but not MMN) are eliminated or at least strongly diminished when the task-irrelevant deviant is predicted by a visual cue (Sussman et al. 2003; Horváth and Bendixen 2012). Cues that provide more specific information about the distracting stimulus are more effective in preventing distraction and the elicitation of P3a and RON (Horváth et al. 2011).

Selective attention related AERPs have been traditionally studied in the context of the classical "cocktail-party" situation described by Cherry (1953). In the simplified dichotic listening model of this situation, participants are exposed to two concurrent messages (one to each ear). Using this paradigm, many studies attempted to decide between the "early" (Treisman 1964; Treisman 1998; Broadbent 1970) versus "late" selection theories of attention (Deutsch and Deutsch 1963; Norman 1968). These theories of attention primarily differ from each other in the placement of a selective filter within the chain of information processing (Broadbent 1958): whereas early selection theories suggest that stimuli are selected for elaborate processing based on simple sensory features (such as pitch) and unattended stimuli do not receive processing beyond extracting these sensory features, late selection theories propose that all stimuli receive elaborate processing and stimuli can therefore be selected on the basis of higher-order properties. (Note that more recent theories of attention do not posit a single selective filter; see e.g., Lavie 1995.) The seminal observation by Hillyard et al. (1973) that selective attention enhances the N1 amplitude for stimuli presented in the to-be-attended channel favoured the early filtering view. However, the findings of Näätanen et al. (1978) of a long-lasting negativity elicited by all attended stimuli, the *Processing Negativity* (PN; Näätänen 1982) challenged this interpretation providing support to late-selection theories. Subsequent studies confirmed both of these effects (Okita 1979; Hansen and Hillyard 1980; Näätänen et al. 1980) and proposed subtraction of the AERP elicited by the non-attended stimuli from that elicited by the attended stimuli as the method to reveal the *Negative Difference* (*Nd*) potential to isolate the AERP correlates of selective attention (Nd; Hansen and Hillyard 1980). The *Nd* is composed of two parts: the early one, termed Nd_e, associated with a gating mechanism preferentially processing the task-relevant stimulus features, and a later part (Nd_l) related to the maintenance of the attentional trace (correspond to the PN). The functional distinction

between the Nd and PN has been debated in detail (Alho et al. 1986a; Alho et al. 1986b; Alho et al. 1994; Teder et al. 1993). Studies showing very early selective attention effects, e.g., at the latency range of the MLR (Woldorff et al. 1987; Woldorff and Hillyard 1991) and possibly even earlier, at the level of the cochlea (Giard et al. 1994) support the interpretation of the Nd_e as a correlate of gating by simple stimulus features. On the other hand, the fact, that the more similar the stimulus to the target the longer the corresponding PN, supports the notion of a comparison with the attentional trace. The frontal scalp distribution of Nd_l (Woods and Clayworth 1987) and the cerebral sources of PN (Giard et al. 1988) are also compatible with the memory-based interpretation of Nd_l. There are several further ERP components related to various facets of attention. However, these are not specific to the auditory modality and thus fall outside the scope of this entry.

AERPs Reflecting Speech and Music processing

The sounds of speech and music may elicit any and all the AERP responses described above. There are, however, also some ERP responses, which arise from events that can be defined in syntactic or semantic terms. It should be noted that most speech-related ERPs can also be elicited through reading. Most AERP responses specific to speech and music have been obtained in paradigms, in which the expectation for the most likely (or simplest) continuation of a sequence of words has been violated. For example, violating the expectation for the first phoneme of the upcoming word elicits a negative shift in the 150–350 ms latency range, termed the *Phonological Mismatch Negativity* (*PMN;* Connolly and Phillips 1994). It is, however, debated, whether this response can be separated from that elicited by words, which are semantically incongruent with respect to the preceding context (D'Arcy et al. 2004; Van den Brink and Hagoort 2004). Violating speech syntax can lead to the elicitation of the *Early Left Anterior Negativity* (*ELAN*) in the 150–200 or the *Left Anterior Negativity* (*LAN*) in the 300–500 ms latency range, depending on the type of violation, whereas potentially correct but syntactically complex sentences elicit the *Syntactic Positive Shift* (*SPS* or *P600*) (for reviews, see Friederici 2002; Hagoort 2008). Violating semantic expectations in speech elicits the *N400* component (Kutas and Federmeier 2011). Musical syntax violations elicit an ELAN-like but predominantly right-hemispheric response, the *Early Right Anterior Negativity* (*ERAN*) in the 180–200 ms or the *Right Anterior-Temporal Negativity* (*RATN*) in the 200–400 ms latency range and N400 has been also be observed in musical models of semantic incongruence (Koelsch and Siebel 2005). For a more detailed discussion of speech- and music-related ERPs, see ► "Electrophysiological Indices of Speech Processing" and ► "Music Processing in the Brain".

Development of AERPs

Previous sections described the AERP responses elicited in adults. Although AERPs can be recorded immediately after birth and even in foetuses within the womb (Draganova et al. 2005), their morphology and functional characteristics widely differ from the adult responses. Further, different AERP components become mature at different times and they often undergo several intermediate phases before reaching adult-like characteristics. As this topic would require a full entry of its own, here we point the reader to some of the existing literature. The most complete reviews of the maturation of AERPs from infancy to adolescence were provided by Wunderlich et al. (2006) and Coch and Gullick (2012). The early infantile development of the AERP components has been summarized by Kushnerenko (2003); for the maturation of the AERPs reflecting auditory change detection, see Jing and Benasich (2006), for large deviations, see Kushnerenko et al. (2013). The maturation of obligatory AERP components from 5 to 20 years of age is covered in Ponton et al. (2000, 2002). AERP maturation during adolescence is described in Bishop et al. (2007). Summarizing these works, one can conclude that the adult AERP morphology characterizes humans from 17/18 years onward and remains more or less unchanged through ageing. There are, however several findings of differences between elderly

and young adults in specific tasks (for a review, see Friedman 2012).

Modelling AERP's: Some General Principles

Theories that seek to explain some of the LLRs have also been explored using more tightly constrained mathematical and computational models. Here we focus on models of the mismatch negativity (MMN) component, as it has arguably received the most widespread attention. Theoretically, MMN has been variously associated with change detection, adaptation, prediction error, novelty detection, and model adjustment, although for some years, there has been controversy as to whether anything more than adaptation is required to explain the experimental data (e.g., see May and Tiitinen 2010 vs. Näätänen et al. 2011).

Using a modelling framework in which exemplars of each of the competing explanations, listed above, were expressed as mathematical functions of stimulation-induced changes in an *unobservable* 'internal state' and resulting *observable* (EEG) responses, Lieder et al. (2013) investigated the ability of each model to explain empirical MMN responses on a trial-by-trial basis. The models were expressed in a rather abstract way, as summarized below, with simple expressions for internal state and response functions (intended to predict stimulus-evoked MMN amplitudes), that captured a range of possibilities for each of the categories. Change detection was modelled with the internal state simply a record of the log frequency of the previous tone in the sequence, and response functions as: a) a flag, set if a difference was detected, b) the signed and c) absolute difference between the frequency of the incoming and previous tone; giving three change detection models. Adaptation was modelled by the exponential decay and recovery of the internal state variable associated with each stimulus frequency, and the response function as a read out of the internal state corresponding to the incoming stimulus. The internal state for the prediction error, novelty detection, and model adjustment accounts was modelled as a Bayesian observer's belief in the tone category of the stimulus, with the evolution of tone category modelled according to a transition matrix derived incrementally from the data according to the 'free-energy-minimisation principle' (Friston 2005). Two prediction error response functions were modelled: prediction errors with respect to sensory input and internal state, respectively. Novelty response functions were modelled as surprise about sensory input and temporal structure (tone category), respectively. Model adjustment response functions were modelled in terms of adjustments to the parameters of the internal model, e.g. mean frequency of a category, expected sequence length, transition probabilities between categories. Simulations showed that, at least at this level of detail, prediction error (with respect to tone category) and model adjustment models (change in expected sequence length, change in transition probabilities between categories), accounted best for the data (Lieder et al. 2013).

On the other hand, May and Tiitinen (2010) have argued strongly that their neural model which includes adaptation on the inputs can explain all MMN data to date; the key mechanism being the activation of fresh afferents by stimuli that deviate in some way from the standards. In this account, MMN is seen as a modulation of the N1 component rather than as a separate component in its own right. The model, consisting of a bank of neural oscillators driven via adapting input synapses, can account for the latency as well as the amplitude of the MMN (May et al. 1999). In addition, extending the model to include local inhibitory feedback circuits, results in a set of non-homogeneous band-pass temporal filters that can also support the topographic representation of stimulus presentation rate (May and Tiitinen 2001). Ringing in these filters is argued to account for the MMN elicited by a missing expected sound. Diverse receptive fields, e.g. to frequency modulations, also allow the model to simulate MMN responses elicited by violations of some abstract rules, such as a repeated tone in a random pattern of ascending tone pairs. However, although adaptation is claimed to be the key to MMN, the model responses also depend upon the amplification of recurrent excitation, lateral inhibition, and the connectivity of the network. The

model thus essentially contains within its changing pattern of adaptation and inhibition, a memory trace of recent activation, and in this sense, contains a memory component embedded within it.

Building on their previous work on a brain-inspired architecture for learning long-term representations of action-perception associations, Garagnani and Pulvermüller (2011) proposed a similar model in which, in addition to adaptation and inhibition, spreading activation through circuits strengthened by learning (long term memory) caused MMN responses to familiar deviants to be larger than that to unfamiliar deviants. They pointed out that only through some form of long term memory mechanism could this differential sensitivity of MMN to familiarity/unfamiliarity be explained. By modelling multiple auditory areas they also provided a novel explanation for differences between the N1 and MMN generators, with N1 being generated in primary auditory areas subject to strong adaptation, and MMN in addition to adaptation also being influenced by reverberating excitation within distributed memory circuits. However, the model processes sequences of static patterns, and as presented, it is not able to account for the sensitivity of MMN to unexpected changes in the timing of sequences, such as the omission MMN (Yabe et al. 1997).

A model that explicitly includes a separate memory module to keep track of the short term history of activation and simulates MMN at a finer level of granularity, i.e. at the level of spiking neurons, was proposed by Wacongne et al. (2012). Memory in the model is implemented using a set of neurons organised into a delay line, i.e. their connectivity ensures that activity passes in one direction across the population, and the progress of activity through this population explicitly represents the timing of the previous event, up to 400 ms. Separate delay lines are used for each tone frequency modelled, thereby also recording their identity. The model simulates MMN by means of prediction errors. Through exposure to tone sequences it learns to generate a prediction of the next tone (both its timing and identity) in a repeating pattern. These predictions are compared with the incoming stimuli in the prediction error units, where mismatches result in a larger signal than matches. The model learns transition probabilities between successive events, as long as they fit within its fixed memory span. In contrast to the adaptation account of MMN, the model relies exclusively on prediction errors. An experiment designed to distinguish between these two explanations for MMN found evidence in favour of a predictive error model of MMN (Wacongne et al. 2012), a result compatible with the findings of Lieder et al. (2013).

A predictive coding account of MMN has also been modelled at a more abstract level using a Kalman filtering (Kalman 1960) approach (Kaya and Elhilali 2013). In this case the timing of events is modelled using a separate filter from the one used to model feature distributions. The advantage of the Kalman filter is that it provides a well-understood way to recursively estimate the system state, refined through analysis of prediction errors, and has been shown to be implementable in the form of a neural network (Szirtes et al. 2005). The model adapts to the variance in observations and, with time, as its predictions improve so its tolerance decreases, making it more sensitive to outliers. Deviants are detected as events not predicted by any existing filter, and trigger the creation of a new set of Kalman filters intended to model a potentially new sound source, making this an interesting framework for more general auditory scene analysis problems, e.g. (Chakrabarty and Elhilali 2013).

In summary, computational models of the theoretical accounts of MMN have begun to be explored. However, so far they have either only been implemented at a rather abstract level; e.g. (Garrido et al. 2009; Lieder et al. 2013; Kaya and Elhilali 2013), focus exclusively on a single mechanism for explaining MMN; e.g. (May and Tiitinen 2001; Wacongne et al. 2012) or account only for MMN responses to unexpected within-event properties (Garagnani and Pulvermüller 2011; Garagnani et al. 2008). The finding, using dynamic causal modelling, that modifications to both feed forward and feedback connections are required (Garrido et al. 2009), and evidence in auditory cortex for adaptation, short term and long term plasticity, recurrent excitation

and inhibition suggests that MMN in the brain may actually depend on the combination of all these factors. Furthermore, while the learning of transition probabilities may be sufficient for some scenarios, in the short term at least, people become sensitive to specific tone patterns; it is unclear whether any of the models discussed here could respond differentially to violations of more extended pattern sequences or more abstract rules.

Utility of AERP for Clinical Practice

Clinical applications of AERPs range from routine practice in audiology, neurotology, neurology, and neurosurgery by ABRs and MLRs (Picton 2010) to highly promising tools for cognitive assessment by some long-latency endogenous components, of which MMN is a prime example. In audiology, ABRs are used universally for hearing screening in neonates failing the Otoacoustic Emission test (OAEs; Robinette and Glatkke 2007). Currently, about 97% of infants are screened for hearing impairment in the USA (Gaffney et al. 2010). ABRs, elicited by click stimuli, are used as a tool for objective audiometry, and ABRs elicited by pure tones can also be used for assessing frequency-specific thresholds in infants (Stapells and Oates 1997; Stapells et al. 1993). In neurotology and neurology, AERPs are combined with the patient's medical history and with an extensive battery of tests for evaluating the anatomy and functional properties of the ear-brain relationship (Picton 2010) in search for an extensive range of disorders of the ear and the auditory pathway, such as Ménière's disease and demyelinating lesions such as Multiple Sclerosis. In these applications, AERPs are used to determine conduction times along the auditory pathway and to localize the anatomical locus of the brain damage with the help of the known origin of the different ABR waveforms (see reviews in Baloh 1997; Chiappa 1997; Lustig et al. 2003). In addition, ABRs are used in combination with evoked potentials from other modalities to monitor coma prognosis (Guérit 2005; Fischer et al. 2006; see below), or in isolation to corroborate brain death (Machado et al. 1991). In the surgical theatre, MLR is used to monitor the depth of anaesthesia in adults (Bell et al. 2004) and children (Kuhnle et al. 2013). It has been recently shown that, compared with the traditional clinical assessment of depth of anaesthesia, MLR monitoring led to a reduction in (a) the amount anaesthetic drug requirement, (b) the use of vasopressors to manage hypotension, and (c) consequential cognitive impairment (Jildenstål et al. 2011).

Regarding cognitive AEPRs, MMN (see above) has shown great promise for potential clinical applications (Näätänen and Escera 2000). Part of this expectation stems from the fact that MMN indexes auditory discrimination accuracy without the requirement to perform some task (i.e., it can be recorded without the patient's collaboration and even in newborn infants; see Alho et al. 1990) and that it can be elicited very reliably, compared with other cognitive event-related potentials (Escera and Grau 1996; Escera et al. 2000b). Yet, after two decades of clinical research (see Näätänen et al. 2012), except for coma monitoring and prognosis no routine clinical application has emerged for the MMN. As for coma monitoring, it has been demonstrated that the presence of MMN in a comatose patient is associated with the return of consciousness (Kane et al. 1993; Fischer et al. 1999), and that as part of a battery of physiological indicators of brain activity, MMN can be used in the decision tree for estimating awakening from coma (Fischer et al. 2006). Given the large variety of disorders and clinical conditions in which impaired MMNs have been observed, it has been suggested that, rather than providing a specific diagnostic measure for any particular disease, the MMN provides an objective index of dysfunction of *N*-metyl-d-aspartate (NMDA) receptor-mediated cognitive functions (Näätänen et al. 2011). In general, due to their high variability and complex functional and anatomical origin, endogenous AERPs can only be employed within large test batteries for diagnostic and monitoring purposes. However, some of these responses provided new insights into the cognitive and emotional aspects of various neurological and psychiatric disorders (e.g., for schizophrenia research using MMN, see Mondragón-Maya et al. 2011).

AERPs: Advantages and Limitations

(A)ERPs provide information about sound-elicited neural activity with millisecond accuracy. Thus they are ideally suited for breaking down the steps of auditory information processing in the brain in the empiricist tradition. It is thus understandable that some of the most recent theoretical developments in the field (e.g., predictive coding theories; Friston 2005) trace back their roots to Helmholtz' (1860/1962) theories of perception. High temporal resolution coupled with the possibility of finding the neural generators of the various ERP responses is also appealing to neurologists and medical doctors, in general. By finding correlations between AERPs and conscious perception on the one hand (such as the link between ORN and the perception of two concurrent sounds; Alain et al. 2001), and discovering the neural mechanisms underlying the observed AERP waveforms on the other hand (e.g., linking SSA and the deviance-detection responses observed in the MLR latency range; Slabu et al. 2010; Grimm et al. 2011; for a review, see Ayala and Malmierca 2013), AERPs can provide a crucial link in understanding the neural mechanisms of perception.

However, there are a number of limitations to the utility of (A)ERPs for research and applications. Firstly, they only reflect a part of the information processing in the brain. When the number of neurons involved in some process is relatively small, or the neurons are distributed over a large area in the brain, or the neural activation is not fully time-locked to the given auditory event, no ERP can be measured. Other methods, such as time-frequency analysis of the EEG, provide better information about these types of processes. AERPs are usually smaller than their visual counterparts. Consequently the signal to noise ratio, where activity not time-locked to the sound onset is regarded as noise, is quite low. This forces one to present many trials to the participant and rely on assumptions which are not fully met by the EEG signal (such as the independence of the signal from the noise, ergodicity, etc.). Further, the accuracy of localizing the sources of neuroelectric activity is limited by the quality of constraints (e.g., anatomical knowledge) required to solve the inverse problem, and the dispersion of the electrical fields. Although magnetoencephalography provides a better spatial resolution, as was already mentioned, AEFs only reflect tangential sources, but not radial ones, thus restricting their general usefulness. In terms of spatial accuracy, other neuroimaging methods, such as fMRI, provide a superior alternative (at the cost of much lower temporal resolution). Further, the correspondence between perception and AERP responses is often not straightforward, as can be gleaned from the often controversial psychological interpretations mentioned in the main text of this entry. Few AERP components can be consistently observed across different stimulus paradigms, thus limiting the validity of most process-based interpretations. Efforts to discover the neural bases of ERP responses must overcome many obstacles. One of the most difficult problems is that whereas individual neurons can mainly be studied in animal models due to the invasive nature of such investigations, it is often difficult to assess how well findings in various species can be extended to characterizing the human brain. Finally, the biggest issue for clinical applications is, as was already mentioned, the large inter- and even intra-individual variability of AERPs.

In summary, AERPs can potentially provide much information about sound processing in the brain, but for extracting this information, better theories and more tightly constrained models, which can integrate information from the diverse fields of anatomy, neuroscience, and psychology, are required.

Cross-References

- ▶ Adaptation in Sensory Cortices, Models of
- ▶ Anatomy and Physiology of the Mammalian Auditory System
- ▶ Auditory Evoked Brainstem Responses
- ▶ Auditory Perceptual Organization
- ▶ Brain Imaging: Overview
- ▶ Electrophysiological Indices of Speech Processing
- ▶ Music Processing in the Brain

References

Alain C, Winkler I (2012) Recording event-related brain potentials: application to study auditory perception. In: Poeppel D, Overath T, Popper AN, Fay RR (eds) The human auditory cortex, Springer handbook of auditory research, vol 43. Springer, New York, pp 69–96

Alain C, Arnott SR, Picton TW (2001) Bottom-up and top-down infl uences on auditory scene analysis: evidence from event-related brain potentials. J Exp Psychol Hum Percept Perform 27:1072–1089

Alho K (1995) Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. Ear Hear 16:38–51

Alho K, Paavilainen P, Reinikainen K, Sams M, Näätänen R (1986a) Separability of different negative components of the event-related potential associated with auditory stimulus processing. Psychophysiology 23: 613–623

Alho K, Sams M, Paavilainen P, Näätänen R (1986b) Small pitch separation and the selective attention effect on the ERP. Psychophysiology 23:189–197

Alho K, Sainio K, Sajaniemi N, Reinikainen K, Näätänen R (1990) Event-related brain potential of human newborns to pitch change of an acoustic stimulus. Electroencephalogr Clin Neurophysiol 77:151–155

Alho K, Teder W, Lavikainen J, Näätänen R (1994) Strongly focused attention and auditory event-related potentials. Biol Psychol 38:73–90

Alho K, Winkler I, Escera C, Huotilainen M, Virtanen J, Jääskeläinen I, Pekkonen E, Ilmoniemi R (1998) Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. Psychophysiology 35:211–224

Alho K, Grimm S, Mateo-León S, Costa-Faidella J, Escera C (2012) Early processing of pitch in the human auditory system. Eur J Neurosci 36:2972–2978

Althen H, Grimm S, Escera C (2011) Fast detection of unexpected sound intensity decrements as revealed by human evoked potentials. PLoS One 6:e28522

Arnott SR, Bardouille T, Ross B, Alain C (2001) Neural generators underlying concurrent sound segregation. Brain Res 1387:116–124

Atienza M, Cantero JL, Dominguez-Marin E (2002) The time course of neural changes underlying auditory perceptual learning. Learn Mem 9:138–150

Ayala YA, Malmierca MS (2013) Stimulus-specific adaptation and deviance detection in the inferior colliculus. Front Neural Circuits 6:89

Baddeley AD, Hitch GJ (1974) Working memory. In: Bower GA (ed) Recent advances in learning and motivation, vol 8. Academic, New York, pp 47–90

Baldeweg T (2006) Repetition effects to sounds: evidence for predictive coding in the auditory system. Trends Cogn Sci 10:93–94

Baldeweg T (2007) ERP repetition effects and mismatch negativity generation: a predictive coding perspective. J Psychophysiol 21:204–213

Baldeweg T, Klugman A, Gruzelier J, Hirsch SR (2004) Mismatch negativity potentials and cognitive impairment in schizophrenia. Schizophr Res 69: 203–217

Baloh RW (1997) Dizziness, hearing loss, and tinnitus. Oxford University Press, New York

Başar E, Rosen B, Başar-Eroglu C, Greitschus F (1987) The associations between 40 Hz-EEG and the middle latency response of the auditory evoked potential. Int J Neurosci 33:103–117

Behroozmand R, Karvelis L, Liu H, Larson CR (2009) Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. Clin Neurophysiol 120: 1303–1312

Bell SL, Smith DC, Allen R, Lutman ME (2004) Recording the middle latency response of the auditory evoked potential as a measure of depth of anaesthesia. A technical note. Br J Anaesth 92:442–445

Bendixen A, Schröger E, Winkler I (2009) I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. J Neurosci 29: 8447–8451

Bendixen A, Jones SJ, Klump G, Winkler I (2010) Probability dependence and functional separation of the object-related and mismatch negativity event-related potential components. NeuroImage 50:285–290

Berti S (2008) Cognitive control after distraction: event-related brain potentials (ERPs) dissociate between different processes of attentional allocation. Psychophysiology 45:608–620

Bishop DVM, Hardiman M, Uwer R, von Suchodoletz W (2007) Maturation of the long-latency auditory ERP: step function changes at start and end of adolescence. Devel Sci 10:565–575

Borgmann C, Ross B, Draganova R, Pantev C (2001) Human auditory middle latency responses: influence of stimulus type and intensity. Hear Res 158:57–64

Bosnyak DJ, Eaton RA, Roberts LE (2004) Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. Cereb Cortex 14: 1088–1099

Bregman AS (1990) Auditory scene analysis. The perceptual organization of sound. MIT Press, Cambridge, MA

Broadbent DE (1958) Perception and communication. Pergamon Press, New York

Broadbent DE (1970) Stimulus set and response set: two kinds of selective attention. In: Mostofsky DI (ed) Attention, contemporary theory and analysis. Appleton Century Crofts, New York, pp 51–60

Busse L, Woldorff MG (2003) The ERP omitted stimulus response to “no-stim” events and its implications for fast-rate event-related fMRI designs. NeuroImage 18: 856–864

Butler BE, Trainor LJ (2012) Sequencing the cortical processing of pitch-evoking stimuli using EEG analysis and source estimation. Front Psychol 3:180

Cebrian AN, Janata P (2010) Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. Brain Res 1342:39–54

Chakrabarty D, Elhilali M (2013) Predictive analysis of two tone stream segregation via extended Kalman filter. In: Proceedings of 2013 47th annual conference information science system (CISS), https://doi.org/10.1109/CISS.2013.6552279

Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25:975–979

Chiappa KH (1997) Evoked potentials in clinical medicine. Lippincott Williams & Wilkins, Philadelphia

Coch M, Gullick MM (2012) Event-related potentials and development. In: Luck SJ, Kappenman ES (eds) The Oxford handbook of event-related potential components. Oxford University Press, New York, pp 475–511

Connolly JF, Phillips NA (1994) Event-related potential components reflect phonological and semantic processing of the terminal words of spoken sentences. J Cogn Neurosci 6:256–266

Cooper RJ, Atkinson RJ, Clark RA, Michie PT (2013) Event-related potentials reveal modelling of auditory repetition in the brain. Int J Psychophysiol 88:74–81

Cornella M, Leung S, Grimm S, Escera C (2012) Detection of simple and pattern regularity violations occurs at different levels of the auditory hierarchy. PLoS One 7: e43604

Costa-Faidella J, Baldeweg T, Grimm S, Escera C (2011a) Interactions between "what" and "when" in the auditory system: temporal predictability enhances repetition suppression. J Neurosci 31:18590–18597

Costa-Faidella J, Grimm S, Slabu LM, Díaz-Santaella F, Escera C (2011b) Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. Psychophysiology 48:774–783

Cowan N (1984) On short and long auditory stores. Psychol Bull 96:341–370

Cowan N (1988) Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. Psychol Bull 104:163–191

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24:87–114

Cowan N, Winkler I, Teder W, Näätänen R (1993) Short- and long-term prerequisites of the mismatch negativity in the auditory event related potential (ERP). J Exp Psychol Learn Mem Cogn 19:909–921

Crowley KE, Colrain IM (2004) A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin Neurophysiol 115:732–744

Curio G (2005) Ultrafast EEG activities. In: Niedermeyer E, Lopes da Silva FH (eds) Electroencephalography. Basic principles, clinical applications, and related fields, 5th edn. Lippincott Williams & Wilkins, Philadelphia, pp 495–504

D'Arcy RCN, Ryner L, Richter W, Service W, Connolly JF (2004) The fan effect in fMRI: left hemisphere specialization in verbal working memory. Neuroreport 15: 1851–1855

Davis PA (1939) Effects of acoustic stimuli on the waking human brain. J Neurophysiol 2:494–499

Davis H, Davis PA, Loomis AL, Harvey EN, Hobart G (1939) Electrical reactions of the human brain to auditory stimulation during sleep. J Neurophysiol 2:500–514

Demany L, Semal C (2007) The role of memory in auditory perception. In: Yost WA, Popper AN, Fay RA (eds) Auditory perception of sound sources, Springer handbook of auditory research, vol 29. Springer, New York, pp 77–113

Deouell LY (2007) The frontal generator of the mismatch negativity revisited. J Psychophysiol 21:188–203

Deouell LY, Bentin S, Giard MH (1998) Mismatch negativity in dichotic listening: evidence for interhemispheric differences and multiple generators. Psychophysiology 35:355–365

Desimone R (1996) Neural mechanisms for visual memory and their role in attention. Proc Natl Acad Sci U S A 93: 13494–13499

Deutsch JA, Deutsch D (1963) Attention: some theoretical considerations. Psychol Rev 70:80–90

Domínguez-Borràs J, Garcia-Garcia M, Escera C (2008) Negative emotional context enhances auditory novelty processing. Neuroreport 19:503–507

Donchin E, Coles MGH (1988) Is the P300 component a manifestation of context updating? Behav Brain Sci 11: 357–374

Draganova R, Eswaran H, Murphy P, Huotilainen M, Lowery C, Preissl H (2005) Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. NeuroImage 28:354–361

Dyson BJ, Alain C, He Y (2005) I've heard it all before: perceptual invariance represented by early cortical auditory-evoked responses. Brain Res Cogn Brain Res 23:457–460

Escera C, Corral MJ (2007) Role of mismatch negativity and novelty-P3 in involuntary auditory attention. J Psychophysiol 21:251–264

Escera C, Grau C (1996) Short-term replicability of the mismatch negativity. Electroencephalogr Clin Neurophysiol 100:549–554

Escera C, Alho K, Winkler I, Näätänen R (1998) Neural mechanisms of involuntary attention to acoustic novelty and change. J Cogn Neurosci 10:590–604

Escera C, Alho K, Schröger E, Winkler I (2000a) Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol Neurootol 5: 151–166

Escera C, Yago E, Polo MD, Grau C (2000b) The individual replicability of mismatch negativity at short and long inter-stimulus intervals. Clin Neurophysiol 111: 546–551

Escera C, Yago E, Alho K (2001) Electrical responses reveal the temporal dynamics of brain events during

A

involuntary attention switching. Eur J Neurosci 14: 877–883
Fabiani M, Gratton G, Federmeier KD (2007) Event-related brain potentials: methods, theory, and applications. In: Cacioppo JT, Tassinary LG, Berntson GG (eds) Handbook of psychophysiology, 3rd edn. Cambridge University Press, Cambridge, pp 85–119
Fischer C, Morlet D, Bouchet P, Luaute J, Jourdan C, Salord F (1999) Mismatch negativity and late auditory evoked potentials in comatose patients. Clin Neurophysiol 110:1601–1610
Fischer C, Luauté J, Némoz C, Morlet D, Kirkorian G, Mauguière F (2006) Improved prediction of awakening or non-awakening from severe anoxic coma using tree-based classification analysis. Crit Care Med 34: 1–5
Folstein JR, Van Petten C (2008) Influence of cognitive control and mismatch on the N2 component of the ERP: a review. Psychophysiology 45:152–170
Friederici AD (2002) Towards a neural basis of auditory sentence processing. Trends Cogn Sci 6:78–84
Friedman D (2012) The components of aging. In: Luck SJ, Kappenman ES (eds) The Oxford handbook of event-related potential components. Oxford University Press, New York, p 512
Friedman D, Cycowicz YM, Gaeta H (2001) The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. Neurosci Biobehav Rev 25:355–373
Friston K (2005) A theory of cortical responses. Phil Trans R Soc Lond Ser B Biol Sci 360:815–836
Friston K, Kiebel S (2009) Cortical circuits for perceptual inference. Neural Netw 22:1093–1104
Gaffney M, Eichwald J, Grosse SD, Mason CA (2010) Identifying infants with hearing loss -United States, 1999–2007. Morb Mortal Wkly Rep 59: 220–223
Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. Proc Natl Acad Sci U S A 78:2643–2647
Garagnani M, Wennekers T, Pulvermüller F (2008) A neuroanatomically grounded Hebbian-learning model of attention-language interactions in the human brain. Eur J Neurosci 27:492–513
Garagnani M, Pulvermüller F (2011) From sounds to words: a neurocomputational model of adaptation, inhibition and memory processes in auditory change detection. Neuroimage 54:170–181. erratum in: Garagnani M, Pulvermüller F (2011) Neuroimage 55: 435–436
Garcia-Larrea L, Lukaszewicz AC, Mauguiere F (1992) Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. Neuropsychology 30:723–741
Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009) The mismatch negativity: a review of underlying mechanisms. Clin Neurophysiol 120:453–463
Giard MH, Perrin F, Pernier J, Peronnet F (1988) Several attention-related waveforms in auditory areas: a topographical study. Electroencephalogr Clin Neurophysiol 69:371–384
Giard MH, Perrin F, Pernier J, Bouchet P (1990) Brain generators implicated in processing of auditory stimulus deviance: a topographic ERP study. Psychophysiology 27:627–640
Giard MH, Collet L, Bouchet P, Pernier J (1994) Auditory selective attention in the human cochlea. Brain Res 633:353–356
Godey B, Schwartz D, de Graaf JB, Chauvel P, Liegeois-Chauvel C (2001) Neuromagnetic source localizationof auditoryevokedfields and intracerebral evoked potentials: a comparison of data in the same patients. Clin Neurophysiol 112:1850–1859
Griffiths TD, Warren JD (2004) What is an auditory object? Nat Rev Neurosci 5:887–892
Grimm S, Escera C, Slabu LM, Costa-Faidella J (2011) Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. Psychophysiology 48:377–384
Grimm S, Escera C (2012) Auditory deviance detection revisited: evidence for a hierarchical novelty system. Int J Psychophysiol 85:88–92
Grimm S, Recasens M, Althen H, Escera C (2012) Ultrafast tracking of sound location changes as revealed by human auditory evoked potentials. Biol Psychol 89: 232–239
Guérit JM (2005) Evoked potentials in severe brain injury. Prog Brain Res 150:415–426
Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25: 5382–5388
Haenschel C, Vernon DJ, Dwivedi P, Gruzelier JH, Baldeweg T (2005) Event-related brain potential correlates of human auditory sensory memory-trace formation. J Neurosci 25:10494–10501
Hagoort P (2008) The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. Phil Trans R Soc Lond B Biol Sci 363: 1055–1069
Handy TC (2005) Event related potentials: a methods handbook. Bradford/MIT Press, Cambridge, MA
Hansen JC, Hillyard SA (1980) Endogenous brain potentials associated with selective auditory attention. Electroencephalogr Clin Neurophysiol 49:277–290
Hansen PC, Kringelbach ML, Salmelin R (eds) (2010) MEG: an introduction to methods. Oxford University Press, New York
Helmholtz H (1860/1962) Handbuch der Physiologischen Optik. Southall JPC (ed) English translation vol 3. Dover, New York
Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. Science 182:177–180
Hölig C, Berti S (2010) To switch or not to switch: brain potential indices of attentional control after task-relevant and task-irrelevant changes of stimulus features. Brain Res 1345:164–175

Horváth J, Bendixen A (2012) Preventing distraction by probabilistic cueing. Int J Psychophysiol 83:342–347
Horváth J, Winkler I (2010) Distraction in a continuous-stimulation detection task. Biol Psychol 83:229–238
Horváth J, Czigler I, Sussman E, Winkler I (2001) Simultaneously active pre-attentive representations of local and global rules for sound sequences. Cogn Brain Res 12:131–144
Horváth J, Czigler I, Winkler I, Teder-Sälejärvi WA (2007) The temporal window of integration in elderly and young adults. Neurobiol Aging 28:964–975
Horváth J, Maess B, Berti S, Schröger E (2008a) Primary motor area contribution to attentional reorienting after distraction. Neuroreport 19:443–446
Horváth J, Winkler I, Bendixen A (2008b) Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? Biol Psychol 79:139–147
Horváth J, Sussman E, Winkler I, Schröger E (2011) Preventing distraction: assessing stimulus-specific and general effects of the predictive cueing of deviant auditory events. Biol Psychol 87:35–48
Itoh K, Yumoto M, Uno A, Kurauchi T, Kaga K (2000) Temporal stream of cortical representation for auditory spatial localization in human hemispheres. Neurosci Lett 292:215–219
Javitt DC, Grochowski S, Shelley AM, Ritter W (1998) Impaired mismatch negativity (MMN) generation in schizophrenia as a function of stimulus deviance, probability, and interstimulus/interdeviant interval. Electroencephalogr Clin Neurophysiol 108:143–153
Jildenstål PK, Hallén JL, Rawal N, Gupta A, Berggren L (2011) Effect of auditory evoked potential-guided anesthesia on consumption of anesthetics and early postoperative cognitive dysfunction: a randomized controlled trial. Eur J Anaesthesiol 28:213–219
Jing H, Benasich AA (2006) Brain responses to tonal changes in the first two years of life. Brain and Development 28:247–256
Jones LA, Baxter RJ (1988) Changes in the auditory middle latency responses during all-night sleep recording. Brit J Audiol 22:279–285
Kalman RE (1960) A new approach to linear filtering and prediction problems. J Basic Eng 82:35–45
Kane NM, Curry SH, Butler SR, Cummings BH (1993) Electrophysiological indicator of awakening from coma. Lancet 341:688
Kaseda Y, Tobimatsu S, Morioka T, Kato M (1991) Auditory middle-latency responses in patients with localized and non-localized lesions of the central nervous system. J Neurol 238:427–432
Kaya E, Elhilali M (2013) A model of auditory deviance detection. In: Proceedings of 2013 47th annual conference on information science and system (CISS), Baltimore, MD, https://doi.org/10.1109/CISS.2013.6552254
Knight RT (1984) Decreased response to novel stimuli after prefrontal lesions in man. Electroencephalogr Clin Neurophysiol 59:9–20
Knight RT (1996) Contribution of human hippocampal region to novelty detection. Nature 383:256–259
Koch D, Sanders LD, Neville HJ (2005) An event-related potential study of selective auditory attention in children and adults. J Cogn Neurosci 17:605–622
Koelsch S, Siebel WA (2005) Towards a neural basis of music perception. Trends Cogn Sci 9:578–584
Kraus N, Ozdamar O, Hier D, Stein L (1982) Auditory middle latency responses (MLRs) in patients with cortical lesions. Electroencephalogr Clin Neurophysiol 54:275–287
Kuhnle GE, Hornuss C, Lenk M, Salam AP, Wiepcke D, Edelmann-Gahr V, Flake G, Daunderer M, Oberhauser M, Müller HH, Feuerecker M (2013) Impact of propofol on mid-latency auditory-evoked potentials in children. Br J Anaesth 110:1001–1009
Kujala T, Tervaniemi M, Schröger E (2007) The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. Biol Psychol 74:1–19
Kushnerenko EV (2003) Maturation of the cortical auditory event-related brain potentials in infancy. Doctoral dissertation, University of Helsinki, Helsinki. https://helda.helsinki.fi/handle/10138/19818
Kushnerenko E, Van den Bergh BRH, Winkler I (2013) Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. Front Psychol 4:595
Kutas M, Federmeier KD (2011) Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Ann Rev Psychol 62:621–647
Lavie N (1995) Perceptual load as a necessary condition for selective attention. J Exp Psychol Hum Percept Perform 21:451–468
Lieder F, Daunizeau J, Garrido MJ, Friston KJ, Stephan KE (2013) Modelling trial-by-trial changes in the mismatch negativity. PLoS Comput Biol 9:e1002911
Lü ZL, Williamson SJ, Kaufman L (1992) Behavioral lifetime of human auditory sensory memory predicted by physiological measures. Science 258:1668–1670
Luck SJ (2005) An introduction to the event -related potential technique. MIT Press, Cambridge, MA
Luck SJ, Kappenman ES (eds) (2012) Oxford handbook of event-related potential components. Oxford University Press, New York
Lustig LR, Niparko J, Minor LB, Zee DS (eds) (2003) Clinical neurotology: diagnosing and managing disorders of hearing, balance and the facial nerve. Martin Dunitz, London
Machado C, Valdés-Sosa P, García-Tigera J, Virues T, Biscay R, Miranda J, Coutin P, Román J, García O (1991) Brain-stem auditory evoked potentials and brain death. Electronecephalogr Clin Neurophysiol 80:392–398
Mäkelä JP, Hari R, Leinonen L (1988) Magnetic responses of the human auditory cortex to noise/square wave transitions. Electroencephalogr Clin Neurophysiol 69:423–430

May P, Tiitinen H (2001) Human cortical processing of auditory events over time. Neuroreport 12:573–577

May PJC, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47:66–122

May P, Tiitinen H, Ilmoniemi RJ, Nyman G, Taylor JG, Näätänen R (1999) Frequency change detection in human auditory cortex. J Comput Neurosci 6: 99–120

Melara RD, Rao A, Tong Y (2002) The duality of selection: excitatory and inhibitory processes in auditory selective attention. J Exp Psychol Hum Percept Perform 28:279–306

Meyer M, Elmer S, Baumann S, Jancke L (2007) Short-term plasticity in the auditory system: differential neural responses to perception and imagery of speech and music. Restor Neurol Neurosci 25:411–431

Mondragón-Maya A, Bernal-Hernández J, Yáñez-Téllez G, Rodríguez-Agudelo Y (2011) Mismatch negativity (MMN) and schizophrenia: a revision. Actas Esp Psiquiatr 39:363–373

Müller MM, Keil A, Kissler J, Gruber T (2001) Suppression of the auditory middle-latency response and evoked gamma-band response in a paired-click paradigm. Exp Brain Res 136:474–479

Munka L, Berti S (2006) Examining task-dependencies of different attentional processes as reflected in the P3a and reorienting negativity component of the human event-related brain potential. Neurosci Lett 396: 177–181

Näätänen R (1982) Processing negativity: an evoked-potential reflection of selective attention. Psychophysiology 24:375–425

Näätänen R (1990) The role of attention in auditory information processing as revealed by event related potentials and other brain measures of cognitive function. Behav Brain Sci 13:201–288

Näätänen R (1992) Attention and brain function. Lawrence Erlbaum Associates, Hillsdale

Näätänen R, Escera C (2000) Mismatch negativity (MMN): clinical and other applications. Audiol Neurootol 5:105–110

Näätänen R, Michie PT (1979) Early selective attention effects on the evoked potential. A critical review and reinterpretation. Biol Psychol 8:81–136

Näätänen R, Picton TW (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

Näätänen R, Winkler I (1999) The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull 125:826–859

Näätänen R, Gaillard AWK, Mäntysalo S (1978) Early selective attention effect on evoked potential reinterpreted. Acta Psychol 42:313–329

Näätänen R, Gaillard AWK, Mäntysalo S (1980) Brain potential correlates of voluntary and involuntary attention. In: Kornhuber AHM, Deecke L (eds) Motivation, motor and sensory processes of the brain: electrical potentials, behavior and clinical use. Elsevier, Amsterdam, pp 343–348

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) 'Primitive intelligence' in the auditory cortex. Trends Neurosci 24:283–288

Näätänen R, Kujala T, Winkler I (2011) Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. Psychophysiology 48:4–22

Näätänen R, Kujala T, Escera C, Baldeweg T, Kreegipuu K, Carlson S, Ponton C (2012) The mismatch negativity (MMN) – a unique window to disturbed central auditory processing in ageing and different clinical conditions. Clin Neurophysiol 123: 424–458

Nagarajan S, Gabriel RA, Herman A (2012) Magnetoencephalography. In: Poeppel D, Overath T, Popper AN, Fay RR (eds) The human auditory cortex, Springer handbook of auditory research, vol 43. Springer, New York, pp 97–128

Ninomiya H, Onitsuka T, Chen CH, Kinukawa N (1997) Possible overlapping potentials of the auditory P50 in humans: factor analysis of middle latency auditory evoked potentials. Electroencephalogr Clin Neurophysiol 104:23–30

Norman DA (1968) Toward a theory of memory and attention. Psychol Rev 75:522–536

Nunez PL, Srinivasan R (2006) Electric fields of the brain: the neurophysics of EEG. Oxford University Press, Oxford

Öhman A (1979) The orienting response, attention and learning: an information processing perspective. In: Kimmel HD, van Olst EH, Orlebeke JF (eds) The orienting reflex in humans. Erlbaum, Hillsdale, pp 443–471

Okita T (1979) Event-related potentials and selective attention to auditory stimuli varying in pitch and localization. Biol Psychol 9:271–284

Pantev C, Hoke M, Lehnertz K, Lütkenhöner B, Anogianakis G, Wittkowski W (1988) Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. Electroencephalogr Clin Neurophysiol 69:160–170

Pantev C, Hoke M, Lehnertz K, Lütkenhöner B (1989a) Neuromagnetic evidence of an ampliotopic organization of the human auditory cortex. Electroencephalogr Clin Neurophysiol 72:225–231

Pantev C, Hoke M, Lütkenhöner B, Lehnertz K (1989b) Tonotopic organization of the auditory cortex: pitch versus frequency representation. Science 246:486–488

Pantev C, Makeig S, Hoke M, Galambos R, Hampson S, Gallen C (1991) Human auditory evoked gamma-band magnetic fields. Proc Natl Acad Sci U S A 88: 8996–9000

Picton TW (2010) Human auditory evoked potentials. Plural Publishing, San Diego

Polich J (2007) Updating P300: an integrative theory of P3a and P3b. Clin Neurophysiol 118:2128–2148

Ponton CW, Eggermont JJ, Kwong B, Don M (2000) Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. Clin Neurophysiol 111:220–236

Ponton C, Eggermont JJ, Khosla D, Kwong B, Don M (2002) Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. Clin Neurophysiol 113: 407–420

Pritchard WS, Shappell SA, Brandt ME (1991) Psychophysiology of N200/N400: a review and classification scheme. In: Jennings JR, Ackles PK (eds) Advances in psychophysiology: a research annual, vol 4. Jessica Kingsley, London, pp 43–106

Rees A, Green GG, Kay RH (1986) Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. Hear Res 23:123–133

Regan D (1989) Human brain electrophysiology. Elsevier, New York

Rinne T, Sarkka A, Degerman A, Schröger E, Alho K (2006) Two separate mechanisms underlie auditory change detection and involuntary control of attention. Brain Res 1077:135–143

Robinette MS, Glatkke TJ (2007) Otoacustic emissions: clinical applications, 3rd edn. Thieme, New York

Salisbury DF (2012) Finding the missing stimulus mismatch negativity (MMN): emitted MMN to violations of an auditory gestalt. Psychophysiology 49:544–548

Sams M, Alho K, Näätänen R (1983) Sequential effects in the ERP in discriminating two stimuli. Biol Psychol 17: 41–58

SanMiguel I, Corral MJ, Escera C (2008) When loading working memory reduces distraction: behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. J Cogn Neurosci 20:1131–1145

SanMiguel I, Widmann A, Bendixen A, Trujillo-Barreto N, Schröger E (2013) Hearing silences: human auditory processing relies on pre-activation of sound-specific brain activity patterns. J Neurosci 33:8633–8639

Schröger E (1996) Neural mechanism for involuntary attention shifts to changes in auditory stimulation. J Cogn Neurosci 8:527–539

Schröger E, Wolff C (1998a) Behavioral and electrophysiological effects of task-irrelevant sound change: a new distraction paradigm. Cogn Brain Res 71:71–87

Schröger E, Wolff C (1998b) Attentional orienting and re-orienting is indicated by human event-related brain potentials. Neuroreport 9:3355–3358

Seki H, Kimura I, Ohnuma A, Saso S, Kogure K (1993) The auditory evoked middle-latency responses (MLRs): their normative variation and generators. Tohoku J Exp Med 170:157–167

Simons RF, Graham FK, Miles MA, Chen X (2001) On the relationship of P3a and the novelty-P3. Biol Psychol 56:207–218

Slabu LM, Escera C, Grimm S, Costa-Faidella J (2010) Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials. Eur J Neurosci 32:859–865

Snyder JS, Alain C, Picton TW (2006) Effects of attention on neuroelectric correlates of auditory stream segregation. J Cogn Neurosci 18:1–13

Sokolov EN (1963) Higher nervous functions: the orienting reflex. Annu Rev Physiol 25:545–580

Sonnadara RR, Alain C, Trainor LJ (2006) Occasional changes in sound location enhance middle latency evoked responses. Brain Res 1076:187–192

Squires NK, Squires KC, Hillyard SA (1975) Two varieties of long-latency positive waves by unpredictable auditory stimuli in man. Electroencephalogr Clin Neurophysiol 38:387–401

Stapells DR, Oates P (1997) Estimation of the pure-tone audiogram by the auditory brainstem response: a review. Audiol Neurootol 2:257–280

Stapells D, Picton T, Durieux-Smith A (1993) Electrophysiologic measures of frequency-specific auditory function. In: Jacobson JT (ed) Principles and applications of auditory evoked potentials. Allyn & Bacon, New York, pp 251–283

Starr A, Don M (1988) Brain potentials evoked by auditory stimuli. In: Picton TW (ed) Handbook of electroencephalography and clinical neurophysiology. Elsevier, Amsterdam

Sussman E (2007) A new view on the MMN and attention debate: the role of context in processing auditory events. J Psychophysiol 21:164–175

Sussman E, Winkler I, Schröger E (2003) Top-down control over involuntary attention-switching in the auditory modality. Psychon Bull Rev 10: 630–637

Szalárdy O, Bőhm T, Bendixen A, Winkler I (2013) Event-related potential correlates of sound organization: early sensory and late cognitive effects. Biol Psychol 93: 97–104

Szirtes G, Páczos B, Lörincz A (2005) Neural Kalman filter. Neurocomputing 65:349–355

Teder W, Alho K, Reinikainen K, Näätänen R (1993) Interstimulus interval and the selective-attention effect on auditory ERPs: "N1 enhancement" versus processing negativity. Psychophysiology 30:71–81

Tong Y, Melara RD, Rao A (2009) P2 enhancement from auditory discrimination training is associated with improved reaction times. Brain Res 1297:80–88

Treisman AM (1964) Selective attention in man. Br Med Bull 20:12–16

Treisman A (1998) Feature binding, attention and object perception. Philos Trans R Soc Lond Ser B Biol Sci 353:1295–1306

Tremblay KL, Kraus N (2002) Auditory training induces asymmetrical changes in cortical neural activity. J Speech Lang Hear Res 45:564–572

Tremblay K, Kraus N, McGee T, Ponton C, Otis B (2001) Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. Ear Hear 22: 79–90

Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. Nat Neurosci 6: 391–398

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24:10440–10453

Van den Brink D, Hagoort P (2004) Influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. J Cogn Neurosci 16:1068–1084

Vanhatalo S, Voipio J, Kaila K (2005) Infraslow EEG activity. In: electroencephalography. Basic principles, clinical applications, and related fields, 5th edn. Lippincott Williams & Wilkins, Philadelphia, pp 489–494

Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci U S A 108:20754–20759

Wacongne C, Changeux JP, Dehaene S (2012) A neuronal model of predictive coding accounting for the mismatch negativity. J Neurosci 32:3665–3678

Winkler I (2007) Interpreting the mismatch negativity. J Psychophysiol 21:147–163

Winkler I, Czigler I (2012) Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. Int J Psychophysiol 83:132–143

Winkler I, Tervaniemi M, Näätänen R (1997) Two separate codes for missing fundamental pitch in the auditory cortex. J Acoust Soc Am 102:1072–1082

Winkler I, Teder-Sälejärvi WA, Horváth J, Näätänen R, Sussman E (2003) Human auditory cortex tracks task-irrelevant sound sources. Neuroreport 14:2053–2056

Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cog Sci 13:532–540

Woldorff MG, Hillyard SA (1991) Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalogr Clin Neurophysiol 79:170–191

Woldorff MG, Hansen JC, Hillyard SA (1987) Evidence for effects of selective attention in the mid-latency range of the human auditory event-related potential. In: Johnson R Jr, Rohrbaugh JW, Parasuraman R (eds) Current trends in event-related potential research. Elsevier, Amsterdam, pp 146–154

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc Natl Acad Sci U S A 90:8722–8726

Woods DL, Clayworth CC (1987) Scalp topography dissociate N1 and Nd components during selective attention. In: Johnson R Jr, Rohrbaugh JW, Parasuraman R (eds) Current trends in event-related potential research. Elsevier, Amsterdam, pp 155–160

Woods DL, Knight RT, Scabini D (1993) Anatomical substrates of auditory selective attention: behavioral and electrophysiological effects of posterior association cortex lesions. Cogn Brain Res 1:227–240

Wu J, Yu Z, Mai X, Wei J, Luo Y (2011) Pitch and loudness information encoded in auditory imagery as revealed by event-related potentials. Psychophysiology 48: 415–419

Wunderlich JL, Cone-Wesson BK, Shepherd R (2006) Maturation of the cortical auditory evoked potential in infants and young children. Hear Res 212: 185–202

Yabe H, Tervaniemi M, Reinikainen K, Näätänen R (1997) Temporal window of integration revealed by MMN to sound omission. Neuroreport 8:1971–1974

Yabe H, Sato Y, Sutoh T, Hiruma T, Shinozaki N, Nashida T, Saito F, Kaneko S (1999) The duration of the integrating windows in auditory sensory memory. Electroencephalogr Clin Neurophysiol Evoked Pot Magn Fields Suppl 49:166–169

Yago E, Escera C, Alho K, Giard MH (2001) Cerebral mechanisms underlying orienting of attention towards auditory frequency changes. Neuroreport 12: 2583–2587

Yago E, Escera C, Alho K, Giard MH, Serra-Grabulosa JM (2003) Spatiotemporal dynamics of the auditory novelty-P3 event-related brain potential. Cogn Brain Res 16:383–390

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C (2001) Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 11:411–423

Yvert B, Fischer C, Bertrand O, Pernier J (2005) Localization of human supratemporal auditory areas from intracerebral auditory evoked potentials using distributed source models. NeuroImage 28:140–153

Zwicker E, Fastl H (1990) Psychoacoustics. Facts and models. Springer, Berlin

Auditory Evoked Brainstem Responses

► Auditory Brainstem Responses

Auditory Evoked Field (AEF)

► Auditory Event-Related Potentials

Auditory Evoked Potential (AEP)

► Auditory Event-Related Potentials

Auditory Frequency-Following Responses

Natàlia Gorina-Careta[1,2,3], Teresa Ribas-Prats[1,2,3], Jordi Costa-Faidella[1,2,3] and Carles Escera[1,2,3,4]
[1]Brainlab-Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain
[2]Institute of Neurosciences, University of Barcelona, Barcelona, Spain
[3]Institut de Recerca Sant Joan de Déu (IRSJD), Barcelona, Spain
[4]Institute for Brain, Cognition and Behavior (IR3C), University of Barcelona, Barcelona, Spain

Synonyms

Complex auditory brainstem response (cABR); Envelope-following response (EFR); Speech auditory brainstem response (sABR)

Definition

The frequency-following response (FFR) is a sustained auditory evoked potential that reflects synchronous neural phase-locking to the spectrotemporal components of the acoustic signal. FFRs are recorded noninvasively from the scalp with electroencephalography (EEG) and with magnetoencephalography (MEG) and emerge at circa 7–15 ms from sound onset to auditory frequencies in the range 100–1500 Hz. By reflecting phase-locked activity to the incoming sounds, FFRs faithfully mimic and are as complex as the eliciting stimulus as it unfolds in time, so that they can be recognized as such when played through a speaker. The FFR has gained recent interest in auditory cognitive neuroscience as it captures with great fidelity the tracking accuracy of the periodic sound features in the ascending auditory system. By decomposing the FFR in the temporal and spectral domains, it is possible to read neural traces from the scalp as sounds are transcribed in neuronal aggregates and how these neural sound traces are shaped by experience, context, and challenging conditions, such as listening in noise, with age and in speech and language disorders.

Detailed Description

Introduction to the FFR

The auditory system is essential for us humans, as it allows to make sense of the sounds around us and to decipher the complex spectrotemporal signals hidden in the ongoing acoustic flow, hence supporting human communication, building our cognitive system through the use of language, and promoting social interaction through the joy of music. Acoustic signals are transduced into a neural code in the inner ear and then released through the auditory nerve to the central nervous system. Understanding how the central auditory system encodes for the different spectrotemporal attributes of sounds and isolates different auditory objects is thus capital to understand the way we use sounds to communicate and interact with others, and the frequency-following response provides a valuable neurophysiological tool to unravel these mysteries. The FFR was originally recorded in animal preparations of the auditory nerve and ascending fibers of the auditory pathway as synchronous phase-locked activity in neuronal aggregates elicited to pure tone stimuli. In humans, it was first recorded by Moushegian, Rupert, and Stillman in 1973 (Moushegian et al. 1973) to pure tone stimuli of frequencies of 0.25, 0.5, 1.0, 1.5, and 2.0 kHz, as they could show compelling evidence of the neuroelectric recordings to follow the cycles of the eliciting stimuli at these different frequencies, and of these neuroelectric signals being of central nervous origin.

The FFR is currently conceived to reflect an aggregation of phase-locked neural activity from multiple generators along the auditory system, although it has been considered for long as a putative measure of subcortical sound encoding. The FFR can be obtained under passive and active listening paradigms, and it is highly sensitive to context-dependent contingencies and to real-time

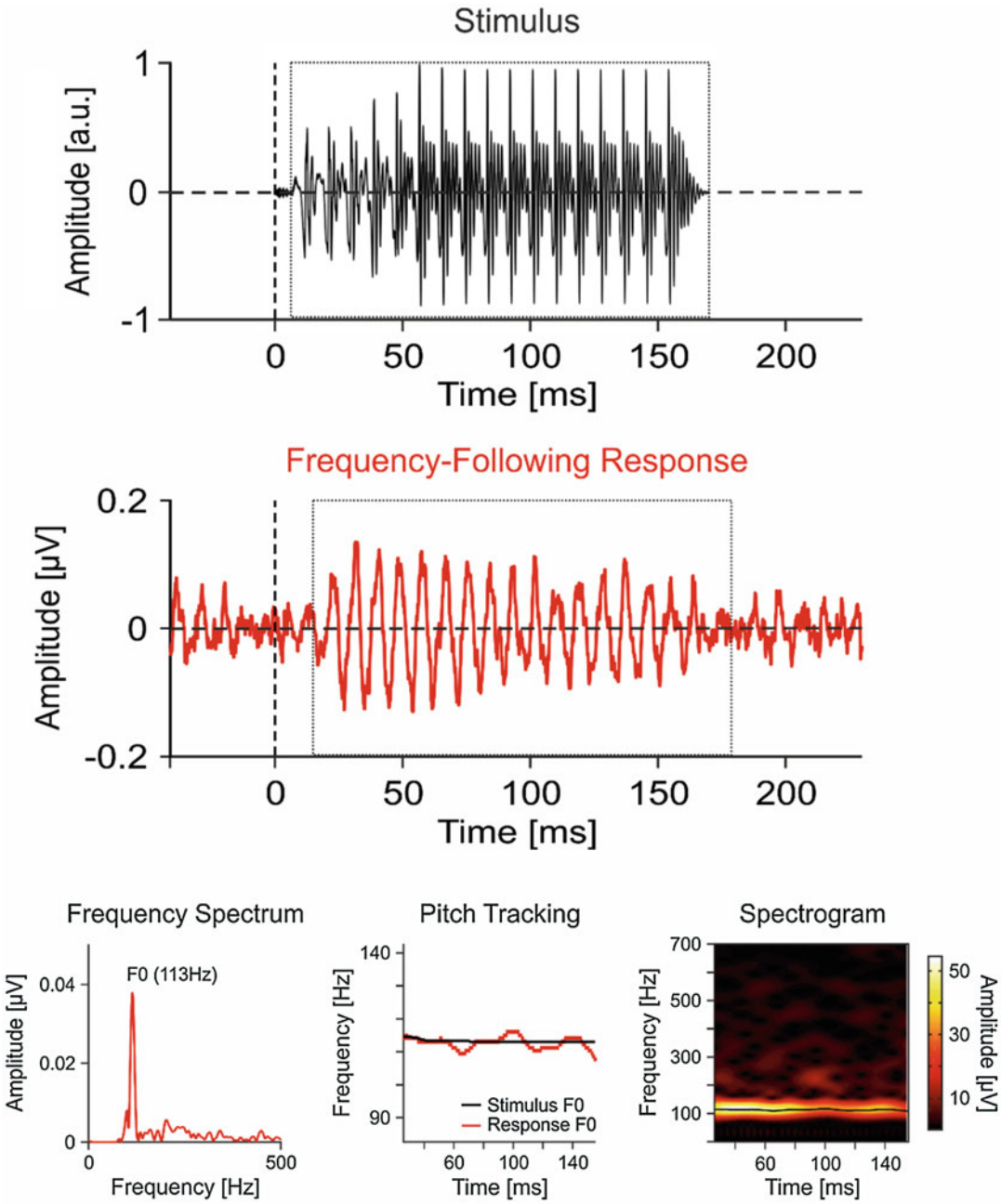
statistical properties of the incoming stimulation (Escera 2017). It is also modulated by short-term auditory training and lifelong auditory experience, including musical training and linguistic competence. Consequently, the FFR has become a major tool in the assessment of the neural encoding of speech sounds in both healthy and clinical populations (Kraus et al. 2017). By means of a range of analytical tools in the temporal and spectral domains, the FFR provides an objective indicator of the fundamental acoustic features intrinsic to speech sounds, including timing (onsets), pitch (fundamental frequency, F0), and timbre (the harmonics information). Specifically, it informs about the latency and amplitude of the auditory input in the time domain and the magnitude of the fundamental frequency and its harmonics in the frequency domain (Kraus et al. 2017) (Fig. 1).

While the term frequency-following response is the most broadly used and probably the most comprehensive one, there are other terms that have been used interchangeably or which highlight a specific aspect or variant in the response. These include *complex auditory brainstem response* (cABR), *speech auditory brainstem response* (sABR), *envelope-following response* (EFR), and *amplitude-modulation following response* (AMFR). As seen, some of these variants refer to the auditory brainstem response (ABR) for good technical reasons: FFRs are recorded with the same equipment and settings as the ABR. In fact, the ABRs can be decomposed into two distinct components, the transient-evoked responses and the FFR (see entry ▶ "Auditory Brainstem Responses"). In addition, as introduced above and explained in further detail below in this entry, the use of the term "brainstem" in naming this auditory evoked potential conveys an anatomical implication that has been overcome by the available evidence. Indeed, FFR was considered since seminal studies to have a primary origin in the inferior colliculus of the auditory midbrain, yet recent studies have indicated that the FFR represents an integrated response of the entire auditory system, with contributions from both subcortical but also cortical centers. Therefore, including anatomical indications on the terminology can lead to misconception. Overall, the scientific community has agreed to the term FFR being the most appropriate one, as it only refers to what the component is: a response that follows the frequency of the incoming stimulus (Kraus et al. 2017).

Importantly, the FFR has a great potential to inform basic and applied questions in learning and communication. Throughout this entry we will review the technical considerations for recording the FFR and computing a large range of FFR-derived measures. We will also discuss its neural origins and how it develops throughout the individual's life-span. Finally, we will characterize the FFR by showing its sensitivity to different auditory contexts and experiences and its role in the study of clinical conditions.

Recording the FFR: Technical Considerations

Selecting appropriate stimuli to elicit reliable FFRs depends on several factors, including specific research purposes as well as taking into account the properties of the electrophysiological response per se. Some of the most commonly used stimuli include *pure tones* (i.e., simple sinusoids), which allow the assessment of responses phase-locked to the sound waveform (response peaks appear with the same periodicity than in the eliciting stimulus); *amplitude-modulated tones*, enabling the study of responses to the amplitude envelope; *complex tones*, used to study both the encoding of different spectral components and the periodic aspects of the sound; *synthetic vowels* and *natural vowels*, used to study the encoding of vowel formant structure with linguistic material; *consonant-vowel syllables*, possibly the best studied, allowing to reveal the encoding of fast formant transitions and sustained responses to vowels (e.g., /da/ vs. /ga/) as well as meaningful pitch contours in tonal languages (slower changes in sound periodicity, e.g., in Mandarin lexical tones); *musical sounds* from different musical instruments, used to study the encoding of harmonic structure and other timbre characteristics; and even *iterated rippled noises*, which are noise stimuli with periodic structure that can be used to study the emergence of perceptual pitch and its neural encoding (Chandrasekaran and Kraus 2010; Skoe and Kraus 2010a).



Auditory Frequency-Following Responses, Fig. 1 Morphology and characteristics of the human scalp-recorded FFR. The frequency-following response (FFR) can be recorded to complex auditory stimuli, such as a consonant-vowel /da/ utterance (upper panel). The middle panel shows the FFR recorded from the Fpz electrode in one single individual (human neonate) to such a stimulus. As can be appreciated, the FFR faithfully mimics the incoming stimulus by phase-locking to its temporal features; notice the same number of cycles within the framed portion in both the stimulus and the neural response. The central origin of the FFR, as opposing to a cochlear origin, is reflected by the short delay in the emergence of phase-locked neuroelectric signal, which corresponds to the neural delay the signal needs to reach the central generating structures within the auditory pathway. In addition to phase-locking to the stimulus temporal components, the FFR also encodes for the spectral features of the incoming stimulus, as can be seen

All these stimulus types are fully periodic or contain periodic sections, and, in order to elicit robust FFRs, their fundamental frequencies (F0) usually range between 80 and 300 Hz, falling within the natural speech range. Their frequency content may span up to 10 kHz, but the FFR signal becomes weaker with increasing frequency (Greenberg 1980), and its phase-locking limits are around 1500 Hz. However, most of the spectral information needed to distinguish vowels is well below 3000 Hz. Stimuli are commonly delivered binaurally at conversational suprathreshold intensities (60–85 dB SPL) to approach naturalistic settings, although monaural stimulation can as well be used. Stimulus duration ranges from several tens of milliseconds to seconds (e.g., 40 ms to 2 s), and interstimulus intervals also vary widely, only limited by experimental time constraints, which may depend on the studied population (recording sessions are shorter in babies or patients suffering from different conditions than in young healthy adults). A common practice is to deliver half of the stimuli with opposite polarity (180° phase shift) to allow averaging and subtracting responses across polarities in order to emphasize responses to the sound envelope (useful to extract pitch tracking measures) or to the spectral content (and study vowel formant structure encoding), respectively (Aiken and Picton 2008; Skoe and Kraus 2010a). Given the small amplitude of the FFR signal (i.e., in the range of a tenth of a μV), a typical recording needs roughly 1000–2000 stimulus presentations to be averaged further per polarity (if applicable).

FFRs are typically recorded with scalp electrodes using essentially the same montage and settings than those used to record ABRs (see entry ▶ "Auditory Brainstem Responses"). In a characteristic montage, an active electrode is placed over the vertex of the scalp or on top of the forehead; a reference electrode is placed on an earlobe (or the mastoid or a high vertebra, at the expense of muscle noise and bone vibration), and a ground electrode is located either on the other earlobe or on the middle of the forehead (Skoe and Kraus 2010a). Impedances are usually kept below 5000 Ω. Recording sampling rate should be at least as high as the double of the highest sound frequency to be studied, although researchers usually oversample (10 or 20 kHz) in order to obtain fine temporal variations in the signal that could be informative (Skoe and Kraus 2010a).

Regarding the most common analysis methods, the FFR is generally computed as an auditory event-related potential (see entry ▶ "Auditory Event-Related Potentials") by averaging or subaveraging per polarity the total number of available trials, in windows lasting typically from −40 ms to 20 ms poststimulus ending. Importantly, the FFR can be distinguished from the cochlear microphonic, a receptor potential generated in the outer hair cells that mimics the incoming stimulus, as the FFR has a 6–10 ms delay poststimulus onset/peaks, while the cochlear microphonic is nearly coincident with the stimulus waveform, or by cancelling the latter via averaging responses to opposite stimulus polarities (Skoe and Kraus 2010a).

Due to the shared characteristics between the FFR and the sound used to elicit it, most of the

Auditory Frequency-Following Responses, Fig. 1 (continued) in the *frequency spectrum*, *pitch tracking* accuracy, and *spectrograms* computed from FFR of this very same individual (bottom panel). The *frequency spectrum* illustrates the amplitude spectral decomposition of the whole FFR, which reveals a clear peak corresponding to the stimulus fundamental frequency. *Pitch tracking* accuracy provides a measure of the ability to track changes in the fundamental frequency along the stimulus entire duration (stimulus frequency depicted in black; response tracking in red). The *spectrogram* provides combined information regarding the frequency and the amplitude at which the FFR is phase-locked to the different spectral components of the incoming stimulus along its entire duration. The color scale from black to white indicates the spectral amplitude in μV, with lighter colors depicting highest amplitude values. Overall, this figure illustrates how the FFR faithfully tracks the eliciting stimulus and its complexity even in a single individual, thus providing a valuable tool to address issues in auditory cognitive neuroscience as well as the assessment of hearing abilities at individual level. Figure adapted and reproduced with permission from Ribas-Prats et al. (2019)

analyses and interpretation on FFRs are based on the acoustic properties of the stimulus, assessing timing and neural synchrony magnitude, as well as phase-locking strength and precision (Skoe and Kraus 2010a). The following are common parameters that are extracted from time and frequency domains, in fixed or sliding windows (to capture changes as a function of time), attempting to disentangle several auditory neural processing features that the FFR may reveal: *stimulus-to-response cross-correlation,* showing the accuracy with which the FFR replicates the stimulus waveform; *neural lag,* an estimation of the temporal delay between stimulus and response; *consistency*, a measure of neural response stability; *root mean square (RMS) amplitude,* indicating the overall magnitude of neural activity over a determined period of time; *pitch strength*, a measure of periodicity based on autocorrelation that reflects the robustness of the response's phase-locking to the stimulus F0 contour; *pitch error,* a measure of how faithfully the FFR encodes the pitch along the stimulus duration, measured as a difference in Hz from the stimulus F0 contour; *spectral amplitude,* usually computed with fast Fourier transform, indicating the magnitude of neural phase-locking at a certain frequency (F0 and harmonics); and *points below noise floor,* describing how the signal can be differentiated from baseline (i.e., pre-stimulus) noise (see Ribas-Prats et al. 2019, for an empirical implementation of all these measures).

Neural Generators

Despite the accumulation of studies on the neural origins of the FFR, no clear picture has emerged so far regarding its anatomical generators, and certain controversy is still being debated. Early seminal studies were addressed to demonstrate that the FFR had a central rather than a cochlear origin, and its sources were attributed to neuronal aggregates in caudal brainstem and midbrain structures, with the inferior colliculus (IC) being the major neuronal source. These seminal studies supported somehow the use of "brainstem" in variants of the nomenclature and its treatment as a putative correlate of subcortical sound encoding. The midbrain origin is supported by the fact that the short latency of the FFR aligns with the latency of the first spikes in IC (Langner and Schreiner 1988) and since FFRs contain phase-locked activity up to 1500 Hz, which spans beyond the upper limit of phase-locking capabilities of cortical neurons (~100 Hz; Aiken and Picton 2008). Additionally, focal lesions (Sohmer et al. 1977) as well as the cryogenic cooling of the IC result in the abolishment of FFRs, with subsequent heating in this later case yielding recovering of FFRs both in the IC and at the scalp (Marsh et al. 1970; Smith et al. 1975). Nevertheless, a mixture of brainstem sources was indeed recognized in the generation of the FFR (Chandrasekaran and Kraus 2010; Tichko and Skoe 2017) and was further supported by other studies that reported weaker contributions of the IC to the FFR, with the major source located in the CN (Gardi et al. 1979a) or in the MGB (Weinberger et al. 1970).

Recently, the controversy around the neural origins of the FFR was renewed with MEG evidence demonstrating that the responses to a complex auditory stimulus of a fundamental frequency close to 100 Hz receive contributions not only from the subcortical nuclei (i.e., the cochlear nucleus and the IC) but also from the medial geniculate body of the thalamus and to a major extent from the auditory cortex (Coffey et al. 2016, 2017). The implications of these findings need, however, a re-examination in the light of the phase-locking capabilities of neuronal aggregates along the auditory hierarchy. Indeed, the upper limit of temporal precision in phase-locked firing reduces with each ascending step in the auditory pathway, so that the ability of neurons to follow fast modulations reduces upstream the auditory hierarchy (Batra et al. 1989; Langner 1992; Joris et al. 2004), and therefore the specific frequency of the eliciting stimulus used to obtain the FFR may play a critical role in engaging multiple sources and a specific configuration of subcortical and cortical generators. Hence, capitalizing on the frequency-specific phase-locking capabilities along the auditory hierarchy, it has been observed that the relative contribution of subcortical and cortical sources to the scalp-recorded FFR varies systematically with stimulus frequency. In fact, the cortical contributions to the scalp-recorded FFR observed were restricted to the lowest

(fundamental) frequencies of the speech spectrum (100 Hz), and recent evidence demonstrated that this cortical contribution to the FFR disappears at frequencies higher than 150 Hz (Bidelman 2018). These findings challenge the assumption of the FFR as a correlate of subcortical sound encoding and support an emerging viewpoint in the literature that the FFR represents an integrated response of the entire auditory system (Kraus and White-Schwoch 2015; Kraus and Slater 2016), with its specific neural origins depending on the frequency of the eliciting sounds.

Developmental Issues: The FFR from Birth to Adulthood

The refinement and modulation of the encoding and representation of complex sounds in the cortico-subcortical auditory system, as reflected by the FFR, remain active and undergo stages of progression along the human life-span. The first FFR recording in human neonates was carried out by Gardi and colleagues in 1979 using low-frequency tone bursts (Gardi et al. 1979b), but it was not until three decades later when the neonates' adultlike capabilities to phase-lock to the incoming stimulus fundamental frequency were fully characterized (Jeng et al. 2010; Ribas-Prats et al. 2019). FFRs recorded in neonates have a similar response morphology to those from adults, in both the phase-locking to the periodicity of the stimulus waveform and its latency, thus confirming that the integrity of the subcortical auditory pathway can be assessed using FFRs from the first day of life at the maternity hospital room.

To understand the encoding and processing of complex sounds in infants, it is important to consider first the development of the auditory pathway during the gestational age. The inner ear and the cochlea become fully mature and functional by 5 months of gestation, which makes the fetus sensory and neurally capable to receive acoustic inputs from both its mother's body and from the external world. However, the only acoustic vibrations that reach the fetus are through the mother's womb, which filters out the input of higher frequencies and allows only the transmission of low-frequency sounds to the fetus. This influence of prenatal listening experience on the neural development of auditory subcortical structures has been supported by studies showing that the FFR amplitude to the fundamental frequency of the eliciting stimuli was clearly observed in neonates (Ribas-Prats et al. 2019) and did not increase significantly in older infants (Anderson et al. 2015). On the other hand, the amplitudes of the FFR to higher frequencies (corresponding to the first formant and higher harmonics) are significantly smaller than the ones of the fundamental frequency (Ribas-Prats et al. 2019) and increase significantly with increasing age (Anderson et al. 2015). The mandatory delay of exposure to high frequencies until birth and the continued myelination of neural structures at the subcortical level during the first year of life may explain the different sensitivity to the input frequency content during the first years of life.

The encoding of spectral and temporal speech components becomes adultlike by the age of 6 months or around, period at which, according to electrophysiological and behavioral studies, the preference for the native language becomes evident. The auditory system suffers rapid maturational changes during the firth months of age, followed by an overshoot during childhood (5–11 year old), and remained stable throughout adulthood until aging-related modulations come into effect (Skoe et al. 2015a).

Contrasting with the abundance of literature describing cortical responses in infants, very few studies have focused on the sound processing at subcortical stages of the auditory pathway in neonates. One remarkable attempt to characterize subcortical auditory processing of complex sounds during the first 3 months of age was by Jeng et al. (2016), who conducted a longitudinal study in which the FFR was recorded in a group of infants in two periods of time: during their first days of life (1–3 days after birth) and at the age of 3 months (Jeng et al. 2016). Although only 13 of the initial group of 44 were recorded at 3 months, the results suggested an improvement in pitch processing with increasing age. These findings are in line with a preliminary work by the same group in which the FFR was collected in one of their participants at 1, 3, 5, 7, and 10 months of age (Jeng et al. 2010). However, further

longitudinal follow-up studies are necessary to fully characterize the developmental stages the FFR may follow.

Effects of Experience-Dependent Plasticity on the FFR

As mentioned above, the FFR has gained recent interest in cognitive auditory neuroscience, as it allows investigation of the biological mechanisms and environmental conditions that modulate the encoding of complex sounds in the auditory hierarchy in service of human communication.

One of the most remarkable features of the FFR is its sensitivity to context-dependent contingencies and to real-time statistical properties of the stimulus. A number of studies have demonstrated that the FFR is able to capture the rapid statistical features of the incoming stimulation, disclosing the encoding of auditory regularities along the auditory hierarchy. In particular, it has been shown that the second harmonic of the FFR is enhanced for both local and global repetitions of a five-tone melody, thus indicating encoding of both global and local statistical regularities within the ongoing stimulation, and that regularity encoding mechanisms might be involved when an auditory object must be separated from background noise (Skoe and Kraus 2010b). In a similar vein, the use of an oddball stimulus sequence with consonant-vowel stimuli revealed that the FFR is not only enhanced for local regularities but also attenuated on its second harmonic in response to a deviant event (e.g., one with low local probability; Slabu et al. 2012). These findings were replicated and extended by subsequent studies, thus supporting the role of the entire auditory hierarchy in extracting statistical information from the acoustic background and demonstrating that context-dependent contingencies and learning-dependent plasticity interact in subcortical stages of the auditory hierarchy (Skoe et al. 2014). In a related account, it was further shown that not only stimulus predictability but also temporal predictability of the incoming stimulation enhances regularity encoding of the acoustic environment, thus indicating that context-dependent contingencies of the ongoing auditory input modulate the encoding of stimulus statistics in the ascending auditory pathway (Gorina-Careta et al. 2016).

The neural sensitivity to stimulus statistics generalizes to more ecological conditions, in which sound patterns were embedded within a single uninterrupted sequence, as was demonstrated in a number of studies in which a series of musical notes were presented in random or patterned sequences. In the patterned sequences, the occurrence of a tone predicted with high accuracy the following one. By using this design, attenuated responses were obtained for the patterned condition compared to the random one, and the more enhanced were the subcortical responses to the patterned condition to the random one, the greater was the individual capability to learn the sequence (Skoe et al. 2013). This sensitivity to stimulus statistics is biased by prior experience and the expectations arising from this experience (Skoe et al. 2015b). Interestingly, the sensitivity to the contingencies of the incoming stimulation is not exclusive of adults but is already seen in children with good reading skills (Chandrasekaran et al. 2009). These results indicate that the whole auditory hierarchy is sensitive to the ongoing stimulus context and that prior experience modulates the responses to the incoming sounds.

Evidence for experience-dependent plasticity has also been provided by the results of short-term training studies, in which FFRs were recorded before and after a period of training (for a review, see Carcagno and Plack 2017). For example, as mentioned above, context-dependent contingencies interact with the effects of short-term training on the accuracy of the encoding of the fundamental frequency (F0) of sounds (Skoe et al. 2014), and, as a result of short-term F0 discrimination training, it has also been observed an improvement of the bilingual robustness of the subcortical temporal encoding (Carcagno and Plack 2011). FFR plasticity has also been investigated after the training on the identification of lexical tones (Song et al. 2008; Chandrasekaran et al. 2012) as well as using general speech-in-noise training protocols (Song et al. 2012). By using this latter one, it has been shown that the subcortical encoding of temporal information

is improved after training. The finding that subcortical auditory processing is not static but can be manipulated by training led to the hypothesis that sensory deficits caused by degraded sound processing could be improved by training. Indeed, it was shown that auditory training can alter the preconscious neural encoding of complex sounds by improving the neural synchrony in the auditory brainstem in children with learning disabilities (Russo et al. 2005).

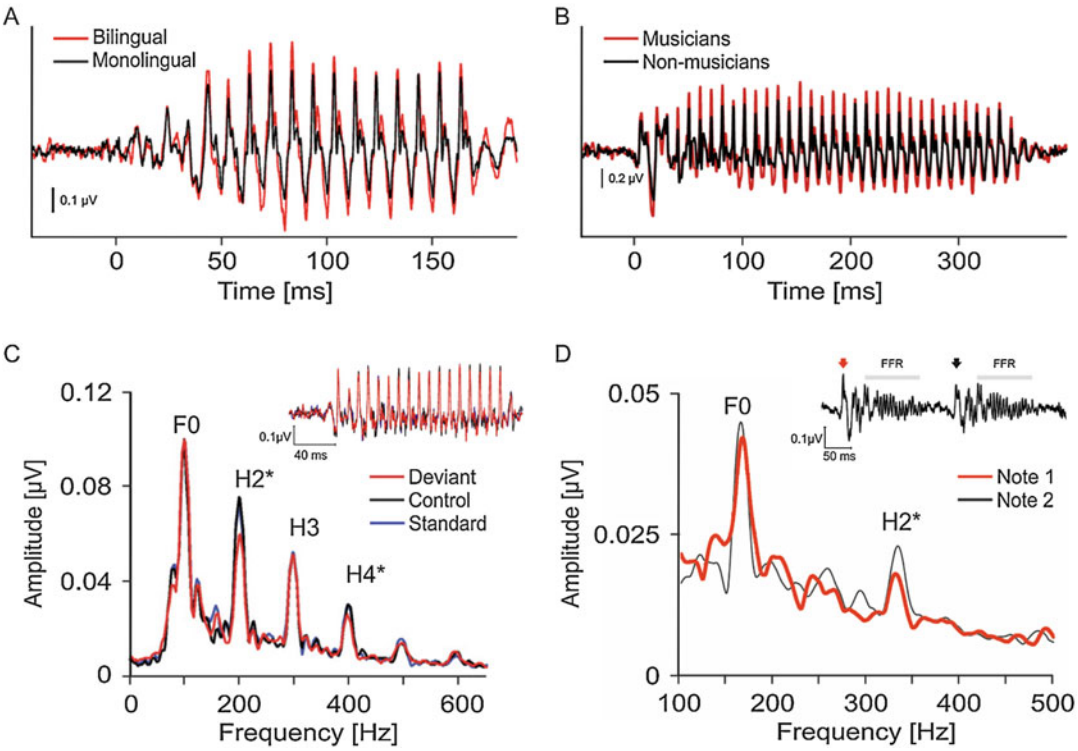
Furthermore, the FFR is not only modulated by short-term auditory training but also by different auditory experiences, such as musical training or language exposure. The first study on the influence of musical training on the FFR was conducted by Musacchia and colleagues (Musacchia et al. 2007), where they demonstrated that musicians have earlier and larger FFRs than nonmusicians to both speech and musical stimuli presented in auditory and audiovisual conditions. Their work was extended by the observation that musicianship enhances the FFR tracking of pitch contours (Wong et al. 2007) and that this experience-dependent plasticity is shaped along the dimensions that are the most behaviorally salient for the listener (Bidelman et al. 2011). Musicians also have a more robust subcortical representation of acoustic stimuli in the presence of noise (Parbery-Clark et al. 2009) and enhanced encoding of speech syllables presented in a predictable condition relative to a variable condition than nonmusicians (Parbery-Clark et al. 2011), thus leading to the hypothesis that subcortical regularity encoding is shaped by musical training and may contribute to the musicians enhanced speech-in-noise perception. Interestingly, the neural changes produced by musical training during childhood are retained in adulthood, as the magnitude of the FFR correlates with how recently the training ceased (Skoe and Kraus 2012).

Language experience is another factor that strongly influences the encoding of speech sounds along the auditory pathway. Bilingual experience enhances the neural responses to the fundamental frequency of sounds (Krizman et al. 2015; Skoe et al. 2017) as well as the subcortical representation of pitch-relevant information (Krizman et al. 2012) and neural consistency, which correlates with both a better attentional control and language proficiency (Krizman et al. 2014). In addition, long-term experience with a tone language (such as Mandarin) sharpens the tuning characteristics of neurons along the pitch axis with enhanced sensitivity to linguistically relevant, rapidly changing sections of pitch contours (Krishnan et al. 2008). In summary, neural encoding of sounds in the subcortical auditory pathway, as revealed by means of the FFR, is shaped by long-term experience with language or music, thus supporting that early sensory processing can undergo experience-dependent plasticity (Fig. 2).

FFR in Challenging Conditions

Human auditory function in challenging conditions can be understood from different perspectives. First, the challenge can be imposed by external situations in which the acoustic signal is degraded by concurrent competing sound streams (i.e., listening in noise) or by room acoustics introducing echoes and reverberation, which necessarily hamper the intelligibility of speech signals. A serious challenge in listening, of an internal nature, is also imposed by hearing loss and even normal aging. A further challenge to central auditory processing is conveyed by clinical conditions, mostly neurodevelopmental, that have been shown to affect auditory, speech, and language perception. In all these domains, the FFR has played a significant role in characterizing the effects of the challenges in the neural encoding of complex, typically speech sounds and its consequences in listening, linguistic, and even cognitive competence.

Both noise- and reverberation-related changes in the FFR are shown in both the time and frequency domains. In particular, noise masks the spectral details of speech, reducing the contrast between the salient frequency features of the sound and the baseline noise (Russo et al. 2004; Li and Jeng 2011). In contrast, reverberation causes a smearing of the spectrotemporal details of the acoustic signal, thus producing a temporal overlap of time-frequency information which results in a systematic degradation in neural periodicity as it degrades the normal phase-locking ability of the FFR (Bidelman and Krishnan 2010). Interestingly,



Auditory Frequency-Following Responses, Fig. 2 Modulation of the FFR by auditory experience and context-dependent contingencies. The FFR to different acoustic stimuli is sensitive and can be modulated by a range of different auditory experiences, such as language exposure and musical training. In particular, this influence is so strong that it can be appreciated in the temporal representation of the FFR when comparing groups with different levels of exposure, for example, to two (bilinguals) instead to only one (monolinguals) language (**a**), or in participants with musical experience as compared to nonmusicians (**b**). The FFR is also highly sensitive to context-dependent contingencies (**c**) and to real-time statistical properties of the stimulus (**d**). In some cases, as the ones illustrated here, differences in the FFRs are not evident in the time domain but can be appreciated in the spectral decomposition of the neural signal. The bottom left panel (**c**) depicts the spectral representation of the FFRs elicited to the same stimulus having three different contextual roles as a function of its probability of occurrence (i.e., as a low-probable [Deviant], red; high-probable [Standard], blue; control, black). Differences became clearly (and statistically) visible in the amplitude of the second harmonic. The bottom right panel (**d**) shows a similar approach to illustrate the effects of short-term statistics (stimulus repetition) on the FFR (first presentation in red; tone repetition in black). Differences between tone statistics (e.g., repetition) are clearly observed in the second harmonic. Figures adapted and reproduced with permission from (**a**) Krizman et al. (2012), (**b**) Skoe and Chandrasekaran (2014), (**c**) Slabu et al. (2012), and (**d**) Skoe and Kraus (2010b)

the two types of acoustic interference do not result in uniform impairment in the speech signal. Indeed, pitch (measured in the encoding of the fundamental frequency) and timbre (assessed by the spectral energy in higher harmonics and spectral envelope) are differentially affected and hence encoded in the FFRs. Little degradation in the FFR fundamental frequency is observed neither with noise nor reverberation, whereas higher harmonics degrade quickly with acoustic interference.

The encoding of speech sounds along the auditory hierarchy with acoustic interference is not stable across the life-span. Rather it naturally declines with age and is impaired in certain auditory disorders. The evoked responses for the three fundamental acoustic features intrinsic to speech sounds are shown to be inefficiently encoded, reduced, or delayed in different ways for distinct clinical populations compared to typically developing controls, but, overall, they all

lead to a weakness in the neural processes that are important for the appropriate auditory processing of the auditory signal. In particular, reduced representation of the fundamental frequency and the harmonics or delayed onset of the FFRs have been observed for children with learning problems (Cunningham et al. 2001) and/or with language deficits and reading disorders such as dyslexia (Banai et al. 2005, 2009; Banai and Ahissar 2006; Chandrasekaran et al. 2009; Anderson et al. 2010; Hornickel et al. 2012; Hornickel and Kraus 2013). Additionally, neural synchrony (timing) and phase-locking (frequency encoding) are also decreased in children with autistic spectrum disorders (Russo et al. 2008, 2009).

Conclusions

In this entry we have attempted to characterize the frequency-following response (FFR) as an auditory evoked potential which is called to play a relevant role in modern auditory cognitive neuroscience. By virtue of its "neurophonic" nature, its ability to reproduce the eliciting sound when played through a speaker, the FFR faithfully reveals how the fine-grained spectrotemporal features of complex auditory stimuli are analyzed and represented along the entire auditory system. Furthermore, by its multidimensional sensitivity to the individual's experience with acoustic surroundings along the lifespan, including long-term exposure to language, music, and noise but also short-term, dynamic processing of auditory streams, the FFR has demonstrated the profound plasticity of the auditory system in the service of auditory perception, the use of language, and communication. By characterizing the complex generating pattern of anatomical structures contributing to the FFR, from the cochlear nucleus to the inferior colliculus, the medial geniculate body of the thalamus, and the auditory cortex, studies with the FFR have revealed the auditory system to be the core of a neural network that interacts with the motor and reward systems to guide and refine our life in sound (Kraus and White-Schwoch 2015). Also, the FFR may have a major translational role in audiologic, neurodevelopmental, and educational domains. In the first place, the FFR recoded at a preschool age has been shown to predict performance across multiple domains of emergent literacy as measured 1 year later (White-Schwoch et al. 2015). Second, given these predictive capabilities of the FFR, one is tempted to speculate that a simple FFR recording at birth (Ribas-Prats et al. 2019) may give a snapshot into the speech-learning abilities on the neonate. Such a test could be routinely applied at the maternity unit, together with the universal newborn hearing screening, right after birth and could lead to early preventive intervention in babies at risk of speech acquisition or developmental delays (Kraus and White-Schwoch 2016). Finally, the FFR may provide a complementary tool in the practice of audiology to objectivize auditory perceptual deficits in patients with complaints about hearing loss in the presence of a normal audiogram and negative testing.

Cross-References

▶ Auditory Brainstem Responses
▶ Auditory Event-Related Potentials

References

Aiken SJ, Picton TW (2008) Envelope and spectral frequency-following responses to vowel sounds. Hear Res 245:35–47

Anderson S, Skoe E, Chandrasekaran B, Kraus N (2010) Neural timing is linked to speech perception in noise. J Neurosci 30:4922–4926

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N (2015) Development of subcortical speech representation in human infants. J Acoust Soc Am 137:3346–3355

Banai K, Ahissar M (2006) Auditory processing deficits in dyslexia: task or stimulus related? Cereb Cortex 16:1718–1728

Banai K, Nicol T, Zecker SG, Kraus N (2005) Brainstem timing: implications for cortical processing and literacy. J Neurosci 25:9850–9857

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Reading and subcortical auditory function. Cereb Cortex 19:2699–2707

Batra R, Kuwada S, Stanford TR (1989) Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. J Neurophysiol 61:257–268

Bidelman GM (2018) Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. NeuroImage 175:56–69

Bidelman GM, Krishnan A (2010) Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Res 1355:112–125

Bidelman GM, Gandour JT, Krishnan A (2011) Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci 23:425–434

Carcagno S, Plack CJ (2011) Subcortical plasticity following perceptual learning in a pitch discrimination task. J Assoc Res Otolaryngol 12:89–100

Carcagno S, Plack CJ (2017) Short-term learning and memory: training and perceptual learning. In: Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN (eds) The frequency-following response. Springer handbook of auditory research, vol 61. Springer, Cham, pp 75–100

Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47:236–246

Chandrasekaran B, Hornickel J, Skoe E, Nicol T, 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

Chandrasekaran B, Kraus N, Wong PCM (2012) Human inferior colliculus activity relates to individual differences in spoken language learning. J Neurophysiol 107:1325–1336

Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ (2016) Cortical contributions to the auditory frequency-following response revealed by MEG. Nat Commun 7:11070

Coffey EBJ, Musacchia G, Zatorre RJ (2017) Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. J Neurosci 37:830–838

Cunningham J, Nicol T, Zecker SG, Bradlow A, Kraus N (2001) Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clin Neurophysiol 112:758–767

Escera C (2017) The role of the auditory brainstem in regularity encoding and deviance detection. In: Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN (eds) The frequency-following response. Springer handbook of auditory research, vol 61. Springer, Cham, pp 101–120

Gardi J, Merzenich M, Mckean C (1979a) Origins of the scalp-recorded frequency-following response in the cat. Int J Audiol 18:353–380

Gardi J, Salamy A, Mendelson T (1979b) Scalp-recorded frequency-following responses in neonates. Int J Audiol 18:494–506

Gorina-Careta N, Zarnowiec K, Costa-Faidella J, Escera C (2016) Timing predictability enhances regularity encoding in the human subcortical auditory pathway. Sci Rep 6:37405

Greenberg S (1980) WPP, No. 52: Temporal neural coding of pitch and vowel quality. UCLA working paper in phonetics

Hornickel J, Kraus N (2013) Unstable representation of sound: a biological marker of dyslexia. J Neurosci 33: 3500–3504

Hornickel J, Anderson S, Skoe E, Yi HG, Kraus N (2012) Subcortical representation of speech fine structure relates to reading ability. Neuroreport 23:6–9

Jeng F-C, Schnabel EA, Dickman BM, Hu J, Li X, Lin C-D, Chung H-K (2010) Early maturation of frequency-following responses to voice pitch in infants with normal hearing. Percept Mot Skills 111:765–784

Jeng F-C, Lin C-D, Wang T-C (2016) Subcortical neural representation to Mandarin pitch contours in American and Chinese newborns. J Acoust Soc Am 139:EL190–EL195

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577

Kraus N, Slater J (2016) Beyond words: how humans communicate through sound. Annu Rev Psychol 67:83–103

Kraus N, White-Schwoch T (2015) Unraveling the biology of auditory learning: a cognitive–sensorimotor–reward framework. Trends Cogn Sci 19:642–654

Kraus N, White-Schwoch T (2016) Newborn hearing screening 2.0. Hear J 69:44

Kraus N, Anderson S, White-Schwoch T (2017) The frequency-following response: a window into human communication (eds: Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN). Springer, Cham

Krishnan A, Swaminathan J, Gandour JT (2008) Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. J Cogn Neurosci 21:1092–1105

Krizman J, Marian V, Shook A, Skoe E, Kraus N (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci 109:7877–7881

Krizman J, Skoe E, Marian V, Kraus N (2014) Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. Brain Lang 128:34–40

Krizman J, Slater J, Skoe E, Marian V, Kraus N (2015) Neural processing of speech in children is influenced by extent of bilingual experience. Neurosci Lett 585:48–53

Langner G (1992) Periodicity coding in the auditory system. Hear Res 60:115–142
Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J Neurophysiol 60:1799–1822
Li X, Jeng F-C (2011) Noise tolerance in human frequency-following responses to voice pitch. J Acoust Soc Am 129:EL21–EL26
Marsh JT, Worden FG, Smith JC (1970) Auditory frequency-following response: neural or artifact? Science 169:1222–1223
Moushegian G, Rupert AL, Stillman RD (1973) Scalp-recorded early responses in man to frequencies in the speech range. Electroencephalogr Clin Neurophysiol 35:665–667
Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci 104:15894–15898
Parbery-Clark A, Skoe E, Kraus N (2009) Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci 29:14100–14107
Parbery-Clark A, Strait DL, Kraus N (2011) Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. Neuropsychologia 49:3338–3345
Ribas-Prats T, Almeida L, Costa-Faidella J, Plana M, Corral MJ, Gómez-Roig MD, Escera C (2019) The frequency-following response (FFR) to speech stimuli: a normative dataset in healthy newborns. Hear Res 371:28–39
Russo N, Nicol T, Musacchia G, Kraus N (2004) Brainstem responses to speech syllables. Clin Neurophysiol 115:2021–2030
Russo N, Nicol T, Zecker SG, Hayes EA, Kraus N (2005) Auditory training improves neural timing in the human brainstem. Behav Brain Res 156:95–103
Russo N, Skoe E, Trommer B, Nicol T, Zecker SG, Bradlow A, Kraus N (2008) Deficient brainstem encoding of pitch in children with autism spectrum disorders. Clin Neurophysiol 119:1720–1731
Russo N, Nicol T, Trommer B, Zecker S, Kraus N (2009) Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Dev Sci 12:557–567
Skoe E, Kraus N (2010a) Auditory brain stem response to complex sounds: a tutorial. Ear Hear 31:302–324
Skoe E, Kraus N (2010b) Hearing it again and again: on-line subcortical plasticity in humans. PLoS One 5:1–9
Skoe E, Kraus N (2012) A little goes a long way: how the adult brain is shaped by musical training in childhood. J Neurosci 32:11507–11510
Skoe E, Krizman J, Spitzer ER, Kraus N (2013) The auditory brainstem is a barometer of rapid auditory learning. Neuroscience 243:104–114
Skoe E, Chandrasekaran B (2014) The layering of auditory experiences in driving experience-dependent subcortical plasticity. Hear Res 311:36–48
Skoe E, Krizman J, Anderson S, Kraus N (2015a) Stability and plasticity of auditory brainstem function across the lifespan. Cereb Cortex 25:1415–1426
Skoe E, Krizman J, Spitzer E, Kraus N (2015b) Prior experience biases subcortical sensitivity to sound patterns. J Cogn Neurosci 27:124–140
Skoe E, Burakiewicz E, Figueiredo M, Hardin M (2017) Basic neural processing of sound in adults is influenced by bilingual experience. Neuroscience 349:278–290
Slabu L, Grimm S, Escera C (2012) Novelty detection in the human auditory brainstem. J Neurosci 32:1447–1452
Smith JC, Marsh JT, Brown WS (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 39:465–472
Sohmer H, Pratt H, Kinarti R (1977) Sources of frequency following responses (FFR) in man. Electroencephalogr Clin Neurophysiol 42:656–664
Song JH, Skoe E, Wong PCM, Kraus N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci 20:1892–1902
Song J, Skoe E, Banai K, Kraus N (2012) Training to improve hearing speech in noise: biological mechanisms. Cereb Cortex 22:1180–1190
Tichko P, Skoe E (2017) Frequency-dependent fine structure in the frequency-following response: the byproduct of multiple generators. Hear Res 348:1–15
Weinberger NM, Kitzes LM, Goodman DA (1970) Some characteristics of the "auditory neurophonic". Experientia 26:46–48
White-Schwoch T, Woodruff Carr K, Thompson EC, Anderson S, Nicol T, Bradlow AR, Zecker SG, Kraus N (2015) Auditory processing in noise: a preschool biomarker for literacy. PLoS Biol 13:1–17
Wong PCM, Skoe E, Russo N, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 10:420–422

Auditory Memory

Dawai Li[1] and Nelson Cowan[2]
[1]Center for Cognitive Neuroscience, Duke University, Durham, NC, USA
[2]Department of Psychological Sciences, University of Missouri-Columbia, Columbia, MO, USA

Synonyms

Acoustic memory; Echoic memory

Definition

Auditory memory is the storage of information about sounds, including both acoustic features (sensory memory) and categorical information about sound categories and multi-sound structure.

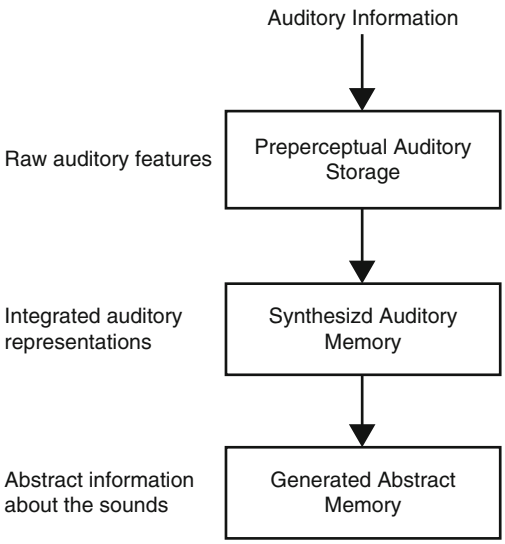
Detailed Description

Auditory memory plays a critical role in various aspects of human activities, such as music, verbal learning, and communication. For example, when a person says, "I said 'rice,' not 'lice,'" the listener must keep the word "rice" in auditory memory to compare it with the word "lice" afterward.

It is widely accepted that auditory memory can be partitioned into three components (Cowan 1984; Crowder 1976; Massaro 1975; Neisser 1967), which Massaro terms preperceptual auditory storage (known also as echoic memory), synthesized auditory memory, and generated abstract memory. Figure 1 illustrates these components and their relationships.

Preperceptual Auditory Storage

Preperceptual auditory storage retains the uncategorized representations of auditory inputs that have not yet been fully processed (Massaro 1975) and is also referred to as short auditory storage (Cowan 1984). It is the auditory counterpart of what is thought of as iconic memory in the visual domain. Preperceptual auditory storage is the first step in auditory processing and starts right after an auditory stimulus enters perception. The duration of preperceptual auditory storage is very short. Most researchers agree that it lasts less than 300 ms. One compelling source of evidence for the duration of preperceptual auditory storage comes from the finding that when a sound is very short (e.g., less than 100 ms), it is still perceived as lasting for about a quarter of a second, which is considered to be the duration of preperceptual auditory storage (for a review see Cowan 1984).



Auditory Memory, Fig. 1 Three phases of auditory memory according to Massaro (1975)

Synthesized Auditory Memory

The auditory features stored in preperceptual auditory storage can be further analyzed to form integrated representations of sound. These integrated representations are considered to be stored in synthesized auditory memory (Massaro 1975). The term "synthesized" refers to the process in which auditory features such as pitch, loudness, and aspects of timbre are analyzed and combined into integrated auditory representations. The duration of the synthesized auditory memory appears to vary from less than 1 s up to 30 s, depending on how it is measured, but it is most often found to be several seconds (Cowan 1984).

The distinction between preperceptual auditory storage and synthesized auditory memory is supported by several lines of research, including backward masking, dichotic listening, and the suffix effect (Cowan 1984). One of the most convincing sources of evidence comes from a backward masking study by Kallman and Massaro (1979). Backward masking refers to the phenomenon that when two sounds are presented sequentially with a very short interval between them, the processing of the first sound (target) sustains interference from the second one (mask). Kallman and Massaro (1979) used two types of sound sequence: (1) standard tone, target tone, and mask (referred to as mask third or "M3") and (2) standard tone, mask, and target tone (referred

to as mask second or "M2"). The participants needed to judge whether the target tone had a higher or lower frequency than the standard tone. In each type of sequence, the interval between the mask and its preceding tone (stimulus onset asynchrony or SOA) was varied, and the mask was either similar to the preceding tone or quite different from it (it was then a white noise). A prediction can be made on the basis of two forms of memory, preperceptual auditory storage and synthesized auditory memory. These two forms can be separately interfered with. In both types of trials, the comparison between the target and standard tones should be impaired by target-mask similarity at a very short SOA because at short SOAs, the similar mask interferes with preperceptual auditory storage of the preceding target tone. Additionally, in the M2 trials only, it is expected that the comparison is always impaired by a similar mask, regardless of the SOA. The reason is that the mask in this procedure comes between the standard and target tones and therefore can interfere with synthesized auditory memory of the standard tone. These expectations exactly match what was found; the target-mask similarity mattered only at short SOAs in the M3 condition, but it mattered at all SOAs in the M2 condition. This finding supports the distinction between preperceptual auditory storage and synthesized auditory memory.

Generated Abstract Memory

The integrated representations in synthesized auditory memory can be further processed to form abstract representations in generated abstract memory (Massaro 1975). The abstract representations are considered to be domain general, meaning that they do not carry information about specific sensory details. Thus, abstract representations generated from each sensory domain (hearing, vision, touch, and so on) are all stored together in the generated abstract memory.

In more recent literature, generated abstract memory is often referred to as "the focus of attention" and is reported to have a core capacity of three to five items when various memory strategies are controlled (Cowan 2001). It is thought that information must be saved in generated abstract memory before high-level thinking about it can occur.

How Auditory Memory Is Used

Although auditory memory is usually partitioned into three phases, all three phases can be used in parallel to process auditory information. Suppose that you are sitting in a noisy airport reading and a stranger asks you what time it is. Even though you did not catch the words immediately, you can still extract the raw auditory information from the preperceptual auditory storage, except for the very last sounds that were masked by someone else nearby talking immediately afterward. The extracted information is then integrated into synthesized auditory memory, which can save the auditory information long enough for you to turn your attention away from the reading and toward the sounds. When your attention is focused on the sounds, you can analyze the sounds based on their memory, using your existing language knowledge. You form a generated abstract memory of what the stranger meant, and you can then respond with the correct time if you have it. This is a typical scenario in which all three phases of auditory memory work together to serve the auditory processing involved in social interactions.

References

Cowan N (1984) On short and long auditory stores. Psychol Bull 96:341–370

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24:87–114

Crowder R (1976) Principles of learning and memory. Erlbaum, Hillsdale

Kallman HJ, Massaro DW (1979) Similarity effects in backward recognition masking. J Exp Psychol 5:110–128

Massaro DW (1975) Experimental psychology and information processing. Rand McNally, Chicago

Neisser U (1967) Cognitive psychology. Appleton, New York

Auditory Pathway

▶ Auditory Processing in Insects

Auditory Perceptual Grouping

► Auditory Perceptual Organization

Auditory Perceptual Organization

Susan Denham[1] and Istvan Winkler[2]
[1]Cognition Institute and School of Psychology, Plymouth University, Plymouth, Devon, UK
[2]Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, MTA, Budapest, Hungary

Synonyms

Auditory perceptual grouping; Auditory scene analysis; Cocktail party problem; Sound source separation

Definition

The process of extracting acoustic features from sound waves and partitioning them into meaningful groups.

Detailed Description

Introduction

Traveling pressure waves (i.e., sounds) are produced by the movements or actions of objects. So sounds primarily convey information about *what is happening* in the environment. In addition, some information about the structure of the environment and the surface features of objects can be extracted by determining how the original (self-generated or exogenous) sounds are filtered or distorted by the environment (e.g., the notion of "acoustic daylight," Fay 2009). In this entry we consider how the auditory systems process sound signals to extract information about the environment and the objects within it.

The auditory system faces a number of specific challenges which need to be considered in any account of perceptual organization: (1) sounds unfold in time; we can't (normally) go back to reexamine them. Therefore, information must be extracted and perceptual decisions made in a timely manner. (2) The information contained within sounds generally requires processing over many timescales in order to extract their meaning (Nelken 2008). For example, a brief impulsive sound may tell the listener that two objects have been in collision, but a series of such sounds is needed in order for the listener to know that someone is clapping rather than walking. (3) Many objects of interest generate sounds intermittently. Therefore, some means for associating temporally discontiguous events are required. (4) Sound pressure waves are additive; what the ear receives is a combination of all concurrently active sound sources and their reflections off any hard surfaces. Many animals and birds communicate acoustically in large social groups, making the problem of source separation particularly tricky (Bee 2012). Despite these challenges, if the auditory system is to provide meaningful information about individual objects in the environment (e.g., potential mates or aggressors), it needs to partition the acoustic features into meaningful groups, a process known as *auditory perceptual organization* or *auditory scene analysis* (Bregman 1990).

Grouping Principles

Auditory Events

Natural environments typically contain many concurrent sound sources, and even isolated sounds can be rather complex, e.g., animal vocalizations contain many different frequency components, and both the frequencies of the components and their amplitudes can vary within a single sound. The problem for the auditory system is to find some way of correctly associating the features which originate from the same sound source. The classical view of this process is that the cochlea decomposes the incoming composite sound waveform into its spectral components, generating a topographically organized array of

signals which sets up the *cochleotopic* (or *tonotopic*) organization found throughout most of the auditory system, up to and including the primary auditory cortex (Zwicker and Fastl 1999). Other low-level features such as onsets, amplitude and frequency modulations, and binaural differences are extracted subcortically and largely independently within each frequency channel (Oertel et al. 2002). These acoustic features are bound together to form auditory *events* (Bertrand and Tallon-Baudry 2000; Zhuo and Yu 2011) or *tokens* (Shamma et al. 2011), i.e., discrete sounds that are localized in time and perceived as originating from a single sound source (Ciocca 2008). Events are subsequently grouped sequentially into patterns, streams, or perceptual objects.

Gestalt Grouping Principles

Perceptual decisions regarding the *causes* of the signals received by the sensors must in general be made with incomplete information (Brunswik 1955). Therefore, potential solutions need to be constrained in some way, e.g., by knowledge about likely sound sources (Bar 2007) or by expectations arising from the recent context (Winkler et al. 2012). In his seminal book, Bregman (1990) pointed out that many such constraints had already been identified by the Gestalt school of psychology (Köhler 1947) early in the twentieth century. The core observation of Gestalt psychology was that individual features form larger perceptual units, which have properties not present in the separate components (von Ehrenfels 1890), and, conversely, that the perception of the components is influenced by the overall perceptual structure (Wertheimer 1912). Focusing primarily on visual stimuli, the Gestalt psychologists described the following grouping principles (laws of perception), here discussed in terms of auditory grouping.

(a) **Good continuation**: Smooth continuous changes in perceptual attributes favor grouping, while abrupt discontinuities are perceived as the start of something new. This principle can operate both within and between individual events.

(b) **Similarity**: Similarity between the perceptual attributes of successive events (e.g., pitch, timbre, location) promotes grouping (Bregman 1990; Moore and Gockel 2002, 2012). Similar to the perception of visual motion (Weiss et al. 2002), it appears that it is not so much the raw difference that is important, but rather the *rate of change*; the slower the rate of change between successive sounds, the more similar they are judged (Winkler et al. 2012). In other words, in the auditory modality, *similarity* and *good continuation* may be equivalent.

(c) **Common fate**: Correlated changes in features promote grouping, recently formalized as *temporal coherence*, i.e., feature correlations within time windows that span periods longer than individual events (Elhilali et al. 2009; Shamma et al. 2011).

(d) **Disjoint allocation (or belongingness)**: Refers to the principle that each element of the sensory input is only assigned to one perceptual object, e.g., exclusive border assignment in Rubin's face-vase illusion. However, although generally true, this principle is sometimes violated in auditory perception, e.g., in *duplex perception*, the same sound component can contribute to the perception of a complex sound as well as being heard separately (Rand 1974; Fowler and Rosenblum 1990).

(e) **Closure**: Objects tend to be perceived as whole even if they are not complete, e.g., a glide continuing through a masking noise if the glide offset is masked (Miller and Licklider 1950; Riecke et al. 2008). This applies more generally to the perception of global patterns (or "Gestalts"), e.g., individual notes are subsumed into a melodic pattern (McDermott and Oxenham 2008) and predictable individual speech sounds are perceived as present even if they are masked or missing (Warren et al. 1988). The auditory system is extraordinarily sensitive to repeating patterns and appears to readily use this cue to parse complex scenes (Winkler 2007; McDermott et al. 2011).

An important concept that emerges from the idea of a "Gestalt" as a pattern is that of predictability. In the case of auditory perception, this refers to expectancies about sound events that have not yet occurred. By detecting patterns (or feature regularities) in the acoustic input, the brain can construct representations that allow it to anticipate or "explain away" (Pearl 1988) future events. In this way Gestalt theory connects to the ideas of unconscious inference (Helmholtz 1885) and perception as hypothesis formation (Gregory 1980).

Auditory Objects

While visual objects are widely accepted as fundamental representational units, the notion of an auditory object is less well established, and there is as yet no universal agreement on how they should be defined, e.g., see Kubovy and Van Valkenburg (2001), Griffiths and Warren (2004), Winkler et al. (2006), Shinn-Cunningham (2008). Based on the Gestalt principles and ideas of perceptual inference, outlined above, Winkler et al. (2009) proposed a definition of an auditory perceptual object as a predictive representation, constructed from feature regularities extracted from the incoming sounds. These object representations are temporally persistent and encode distributions over featural and temporal patterns, determined by the current context. The consolidated object representation therefore refers to *patterns of sound events*; individual sound events are processed within the context of the whole to which they belong. This definition of an auditory perceptual object is compatible with the definition of an auditory stream, as a coherent sequence of sounds separable from other concurrent or intermittent sounds (Bregman 1990). However, whereas the term "auditory stream" refers to a phenomenological unit of sound organization, with separability as its primary property, the definition proposed by Winkler et al. (2009) emphasizes the extraction and representation of the unit as a pattern with predictable components (Winkler et al. 2012). While the usage of the term object is not universally accepted within the auditory domain, we will use it in this entry as defined by Winkler et al. (2009).

Auditory Scene Analysis

In order to determine the perceptual qualities of individual sound events, the brain must first bind their component features even though the number of concurrent auditory objects and which features belong to each is unknown a priori; this must be inferred incrementally from the ongoing sensory input. Therefore, it is clear that the auditory system needs to use (top-down) contextual information to guide its grouping decisions and some means for evaluating these decisions and revising them in the event that they prove to be incorrect. In the currently most widely accepted framework describing perceptual sound organization, auditory scene analysis, Bregman (1990) proposes two separable processing stages. The first stage is suggested to be concerned with partitioning sound events into potential groups based primarily on featural similarities and differences. The second stage, within which prior knowledge and task demands exert their influence, is a competitive process between candidate organizations that determines which one is perceived. Within this framework there are two types of grouping: *simultaneous* grouping based on concurrent cues and *sequential* grouping based on contextual temporal cues. For the reasons outlined above, these two are not really distinct (simultaneous cues are influenced by prior sequential grouping, e.g., Darwin et al. (1995) and Bendixen et al. (2010b), just as sequential grouping is influenced by the perceptual qualities of individual events (simultaneous grouping) (Bregman 1990); nevertheless, they provide a useful starting point for models of auditory scene analysis.

Simultaneous Grouping

In the absence of sequential grouping cues, there are some features which automatically trigger the formation of individual sound events; for reviews see Darwin and Carlyon (1995) and Ciocca (2008). Common onsets and offsets form clear temporal boundaries, and the strategy adopted by the auditory system is to match onsets to offsets (including similarities between features and temporal proximity) in order to segregate perceptual events (Nakajima et al. 2000). Harmonicity

(i.e., the presence of frequency components which are integer multiples of a common fundamental frequency) is another important grouping cue (Darwin and Carlyon 1995). For example, when one component of a complex harmonic tone is mistuned, listeners perceive two concurrent sounds, a complex tone consisting of the harmonically related components and a pure tone, corresponding to the mistuned component (Moore et al. 1986). However, not all acoustic features trigger concurrent grouping, e.g., a location cue (common interaural time differences) between a subset of frequency components within a single sound event does not generate a similar segregation of component subsets within individual sound events (Culling and Summerfield 1995).

Another important strategy for segregating sound events is template matching. If people have prior knowledge of events, then it is possible to hear them out. This effect was exploited in the many double-vowel experiments used to test the influence of different acoustic features, e.g., Assmann and Summerfield (1990) and Summerfield and Assmann (1991), and even in the absence of featural differences, it was shown that known vowel sounds can be identified well above chance (Assmann and Summerfield 1989). This template-matching phenomenon appears to be rather general and applies to any sound that is repeated. The auditory system is very sensitive to repetition (Teki et al. 2011). If a previously unheard sound is repeated against a different background, then it can be segregated and identified significantly above chance, even with only a single repetition, and even if many of usual grouping cues are absent (McDermott et al. 2011). Similarly, arbitrary repeated noise segments can be rapidly learnt within a few trials (Agus et al. 2010).

Models of Event Formation

Many models have been developed to investigate simultaneous grouping and the segregation of perceptual events, e.g., see models described in Wang and Brown (2006). A model of auditory saliency which used low-level cues of spectral and temporal contrast to highlight salient events in continuous noisy soundscapes predicted human event detection very well (Kayser et al. 2005). Temporal contrasts effectively highlight onsets and offsets, while spectral peaks carry information about the resonances of sound sources and to some extent their identity (von Kriegstein et al. 2007). The segregation of overlapping events using pitch cues has been widely explored (c.f. Pitch Perception, Models), e.g., for explaining enhanced double-vowel segregation (de Cheveigne et al. 1995). The segregation of events using repetition was shown to be possible in principle by using a combination of cross-correlation and averaging to incrementally build a representation of the repeated target (McDermott et al. 2011). Because of the importance of longer-term context on grouping, none of these models provide general solutions to the problem of auditory scene analysis; nevertheless, they provide important building blocks in this process.

Sequential Grouping

Sequential grouping generally conforms to the Gestalt principles of *similarity*/*good continuation* and *common fate*. In contrast to concurrent grouping, sequential grouping is necessarily based on some representation of the preceding sounds; for reviews, see (Moore and Gockel 2002; Carlyon 2004; Haykin and Chen 2005; Snyder and Alain 2007; Ciocca 2008; Shamma and Micheyl 2010; Shamma et al. 2011; Moore and Gockel 2012). Most studies of this class of grouping have used sequences of discrete sound events to investigate the influences of acoustic features and temporal structure. In the most widely used experimental approach (termed the auditory streaming paradigm), sequences of alternating sound events differing in some feature(s) are presented to listeners (van Noorden 1975). When the feature separation is small and/or they are delivered at a slow pace, listeners predominantly hear a single *integrated* stream containing all the sounds. With large feature separation and/or fast presentation rates, listeners report hearing the sequence separate out into two *segregated* streams. In this there is a cue trade-off: smaller feature differences can be compensated with higher presentation rates and

vice versa (van Noorden 1975). Differences in various auditory features, including frequency, pitch, loudness, location, timbre, and amplitude modulation, have been shown to support auditory stream segregation (Vliegen and Oxenham 1999; Grimault et al. 2002; Roberts et al. 2002). Thus it appears that sequential grouping is based on perceptual similarity, rather than on specific low-level auditory features (Moore and Gockel 2002, 2012). Temporal structure has also been suggested as a key factor in segregating streams either by guiding attentive grouping processes (Jones 1976; Jones et al. 1981; Large and Jones 1999) or through temporal coherence that binds correlated component features in the auditory input (Elhilali et al. 2009; Shamma and Micheyl 2010; Shamma et al. 2011, 2013).

Models of Auditory Streaming

Early models of auditory streaming, e.g., Beauvois and Meddis (1991), focused on the relationship between frequency differences and event rate and the proposal that streaming could be explained almost exclusively by peripheral channeling mechanisms (Hartmann and Johnson 1991) or the degree of overlap between neural responses to each of the alternating tones, e.g., McCabe and Denham (1997). In these models the perceptual decision was represented by levels of activation across a spatial array of neurons; see also Micheyl et al. (2005) for a similar interpretation of neural activity in primary auditory cortex. A different approach in which grouping is signaled by temporal correlations within network responses was proposed by Wang, Brown, and colleagues (Brown and Wang 2006; Wang and Chang 2008). For example, the model proposed by Wang and Chang (2008) consists of a 2-dimensional array of oscillators with one dimension representing frequency and the other external time. Units are connected by local excitatory connections and by global inhibition. Characteristic results of classical auditory streaming experiments (van Noorden 1975) are simulated by including strong local excitatory connections (encouraging synchronization) and weaker long-range connections (which are easily overcome by inhibition and therefore encourage desynchronization). Sensitivity to event rate is modeled by dynamic weight adjustments. However, while the representation of grouping is different from the models previously outlined, this model also depends on peripheral channeling and the degree of overlap in the incoming activity patterns to determine its grouping decision.

A similar focus on temporal coherence (in this case the average correlation within a sliding window 50–500 ms in duration) is seen in the model of streaming proposed by Elhilali and colleagues, e.g., Elhilali and Shamma (2008) and Shamma et al. (2011) (Note, Figs. 6 and 9 in this entry have incorrect colour scale labels (0% and 100%, interchanged; Shamma and Elhilali (2013)). The computational model developed by Elhilali and Shamma (2008) extracts multiple features from the incoming acoustic input including frequency, pitch, direction, and spectral shape and assigns the resulting activity patterns to one of two clusters which come to represent the properties of the events in each stream. The temporal coherence measure is used to determine which components should be grouped. The clusters compete to incorporate each event, and the winning cluster uses the event features (as determined by the grouping process) to refine its representation. These correlation-based models overcome a problem faced by the population separation account of streaming (Micheyl et al. 2005) that predicted widely separated components would be segregated even if they overlapped in time, which is not the case (Elhilali et al. 2009). They also provide a means for binding the component features of an event, not considered in the earlier models. Later refinements to the temporal coherence account of streaming (Shamma et al. 2011, 2013), included the strong claims that (a) feature binding occurs only with attention, i.e., attention is responsible for grouping features that belong to the foreground object, c.f. (Treisman 1998), and (b) all other features remain ungrouped in an undifferentiated background. However, the proposed role of attention in feature binding has long been debated in the visual domain, e.g., Duncan and Humphreys (1989), and it is not consistent with the results of experiments testing feature binding in the absence of attention by

recording auditory event-related potentials (AERP) in response to rare feature combinations (Takegata et al. 2005; Winkler et al. 2005a).

Competition and Selection

The models described above all conform to the assumptions that in response to alternating two-tone sequences, (a) auditory perception always starts from the integrated organization and (b) that eventually a stable final perceptual decision is reached (Bregman 1990). However, it has been found, when listeners report their percepts continuously while listening to such sequences for long periods, that perception fluctuates between different perceptual organizations (Winkler et al. 2005b; Pressnitzer and Hupe 2006). Perceptual switching occurs in all listeners and for all combinations of stimulus parameters tested (Anstis and Saida 1985; Roberts et al. 2002; Denham and Winkler 2006; Pressnitzer and Hupe 2006; Schadwinkel and Gutschalk 2011; Denham et al. 2012), even combinations very far from the ambiguous region identified by van Noorden (1975). Furthermore, for stimuli with parameters that strongly promote segregation, participants often report hearing segregation first (Deike et al. 2012; Denham et al. 2012). It has also been found that perceptual organizations other than the classic integrated and segregated categories may be reported (Bendixen et al. 2010a, 2012; Bőhm et al. 2012; Denham et al. 2012; Szalárdy et al. 2012), showing that auditory perceptual organization in response to alternating two-tone sequences is multistable (Schwartz et al. 2012).

The notion of perceptual multistability is challenged by everyday subjective experience of a world perceived as stable and continuous and by experimental results obtained by averaging over the reports of different listeners, which generally show that within the initial 5–15 s of two-tone sequence, the probability of reporting segregation monotonically increases (termed the buildup of auditory streaming) (but see Deike et al. (2012)). For these reasons it has been suggested that perceptual multistability observed in the auditory streaming paradigm may be simply a consequence of the artificial stimulation protocol used. However, there is a growing body of experimental data supporting the existence of multistability and just as visual multistability has provided new insights into visual processing, e.g., Kovacs et al. (1996); it seems likely that understanding spontaneous changes in the perception of unchanging sound sequences will help throw new light on auditory perception.

Modeling Multistability in Auditory Streaming

Multistability of auditory perceptual organization cannot be explained by any of the theories or models outlined above, which all have essentially one fixed attractor. Models of visual multistability have a longer history, e.g., Laing and Chow (2002); Shpiro et al. (2009); van Ee (2009). These models typically contain three essential components (Leopold and Logothetis 1999): (a) mutual inhibition between competing stimuli to ensure *exclusivity* (i.e., perceptual awareness generally switches between the different alternatives rather than fusing them), (b) adaptation to ensure the observed *inevitability* of perceptual switching (the dominant percept cannot remain dominant forever), and (c) noise to account for the observed *stochasticity* of perceptual switching (successive phase durations are largely uncorrelated, and the distribution of phase durations resembles a gamma or log-normal distribution) (Levelt 1968). The questions for auditory multistability are what are the competing entities, and what form does this competition take in order to explain dynamic nature of perceptual awareness reported by listeners.

The computational model of auditory multistability proposed by Mill et al. (2013) is based on the idea that auditory perceptual organization rests on the discovery of recurring patterns embedded within the stimulus, constructed by forming associations (links) between incoming sound events and recognizing when a previously discovered sequence recurs and can thus be used to predict future events. These predictive representations, or *proto-objects* (Rensink 2000; Winkler et al. 2012), compete for dominance with any other proto-objects which predict the same event (a form of local competition) and are the candidate set of representations that have the potential to become the perceptual objects of

conscious awareness. This model accounts for the emergence of, and switching between, alternative organizations; the influence of stimulus parameters on perceptual dominance, switching rate, and perceptual phase durations; and the buildup of auditory streaming. In a new sound scene, the proto-object that is the easiest to discover determines the initial percept. Since the time needed for discovering a proto-object depends largely on the stimulus parameters (i.e., to what extent successive sound events satisfy/violate the *similarity/good continuation* principle), the first percept strongly depends on stimulus parameters. However, the duration of the first perceptual phase is independent of the percept (Hupe and Pressnitzer 2012), since it depends on how long it takes for *other* proto-objects to be discovered (Winkler et al. 2012). The model also accounts for the different influences of similarity and closure on perception; the rate of perceptual change (similarity/good continuation) determines how easy it is to form the links between the events that make up a proto-object, while predictability (closure) does not affect the discovery of proto-objects, but can increase the competitiveness (salience) of a proto-object once it has been discovered (Bendixen et al. 2010a).

Neural Correlates of Perceptual Organization

Neural responses to individual sounds are profoundly influenced by the context in which they appear (Bar-Yosef et al. 2002). The question is to what extent the contextual influences on neural responses reflect the current state of perceptual organization. This question has been addressed by a number of studies ranging in focus from the single neuron level (c.f. stimulus-specific adaptation) to large-scale brain responses (c.f. auditory evoked potentials), and the results provide important clues about the processing strategies adopted by the auditory system.

Studies investigating single neuron responses to alternating tone sequences, e.g., Fishman et al. (2004), Bee and Klump (2005), Micheyl et al. (2005)), and Micheyl et al. (2007), have shown an effect called *differential suppression*, i.e., at the start of the sequence, the neuron responds to both tones, but with time the response to one of the tones (typically corresponding to the best frequency of the cell) remains relatively strong, while the response to the other tone diminishes. Since neuronal sensitivity to frequency difference and presentation rate was found to be consistent with the classical van Noorden (1975) parameter space, it was claimed that differential suppression was a neural correlate of perceptual segregation (Fishman et al. 2004). This was supported by the finding that spike counts from neurons in primary auditory cortex predict an initial integration/segregation decision closely matching human perception (Micheyl et al. 2005; Bee et al. 2010). However, differential suppression does not account for perceptual multistability or for the perception of overlapping tone sequences (Elhilali et al. 2009); therefore, while differential suppression may be a necessary component of the auditory streaming process, it does not provide a complete explanation.

Auditory event-related brain potentials (AERPs) represent the synchronized activity of large neuronal populations, time locked to some auditory event. Because they can be recorded noninvasively from the human scalp, they have been widely used to study the brain responses accompanying auditory stream segregation; c.f. auditory event-related potentials, especially long-latency AERP responses. Three AERP components are of particularly relevance in this regard: (a) the "object-related negativity" (ORN) which signals the automatic segregation of concurrent auditory objects (Alain et al. 2002), (b) the amplitude of the auditory P1 and N1 which varies depending on whether the same sounds are perceived as part of an integrated or segregated organization (Gutschalk et al. 2005; Szalárdy et al. 2013), and c) the mismatch negativity (MMN; Näätänen et al. 1978) which has been used as an indirect index of auditory stream segregation, e.g., Sussman et al. (1999); Nager et al. (2003); Winkler et al. (2003a); Gutschalk et al. (2005).

The detection and representation of regularities by the brain, as indexed by the MMN, provided the basis for the definition of an auditory object

proposed by Winkler et al. (2009). Using evidence from a series of MMN studies, they defined an auditory object as a perceptual representation of a possible sound source, derived from regularities in the sensory input (Winkler 2007, 2010) that has temporal persistence (Winkler and Cowan 2005) and can link events separated in time (Näätänen and Winkler 1999). This representation forms a separable unit (Winkler et al. 2006) that generalizes across natural variations in the sounds (Winkler et al. 2003b) and generates expectations of parts of the object not yet available (Bendixen et al. 2009).

It should be pointed out that while traditional psychological accounts of auditory perceptual organization implicitly or explicitly refer to representations of objects, there are models of auditory perception which are not concerned with positing a representation directly corresponding auditory objects. The hierarchical predictive coding model of perception, e.g., Friston and Kiebel (2009), includes predictive memory representations, which are in many ways compatible with the notion of auditory object representations (Winkler and Czigler 2012), but no explicit connection with object representations is made. Shamma and colleagues' temporal coherence model of auditory stream segregation (Elhilali and Shamma 2008; Elhilali et al. 2009; Shamma et al. 2011, 2013) provides another way to avoid the assumption that object representations are necessary for determining sound organization; instead it is proposed that objects are essentially whatever occupies the perceptual foreground and exist only insofar as they *do* occupy the foreground. In summary, there is currently little consensus on the role of auditory object representations in perceptual organization, and the importance placed on object representations by the various models and theories differs markedly.

fMRI studies of auditory streaming have found neural correlates in a number of brain regions. In one of the earliest studies, Cusack (2005) failed to find differential activity in auditory cortex corresponding to perceptual organization into one or two streams, but he did find such activity in the intraparietal sulcus, an area associated with cross-modal processing and object numerosity. Shortly afterwards Wilson et al. (2007) showed that auditory cortical activity increased with increasing frequency difference and that as the frequency difference increased, the cortical response changed from being rather phasic (i.e., far stronger at the onset of the sequence) towards a more sustained response throughout the stimulus sequence. Taking a closer look at the dynamics of cortical activity associated with perceptual switching, Kondo and Kashino (2009) showed that both auditory cortex and thalamus are involved, with an increase in thalamic activity preceding that in cortex associated with a switch from the nondominant to the dominant percept and, conversely, an increase in cortical activity preceding that in thalamus associated with a switch from the dominant to the nondominant percept. They also found differential activation in posterior insular cortex and in the cerebellum. Interestingly, activations in the cerebellum and thalamus are negatively correlated in auditory streaming, with the left cerebellar activation level increasing with the rate of perceptual switching and thalamus (medial geniculate) decreasing (Kashino and Kondo 2012). Consistent with these findings, Schadwinkel and Gutschalk (2011), using a different stimulus paradigm which allowed them to influence the timing of perceptual switching, found transient auditory cortical activation associated with perceptual switching and a further transient activation in inferior colliculus, although whether the inferior colliculus is responsible for triggering switching or simply reflects the transient switching activation in cortex is not clear. In summary, neural correlates of auditory streaming have been found in many areas within the auditory system and beyond, suggesting that creating and switching between alternative perceptual organizations involve a broadly distributed network within the brain.

Conclusions and Open Questions

The Gestalt principles and their application to auditory perception instantiated in Bregman's

(1990) two-stage auditory scene analysis framework provided the initial basis for understanding auditory perceptual organization, and recent proposals have extended this framework in interesting ways. Nevertheless, there remain many unanswered questions and there have been few, if any, attempts to build neuro-computational models capable of dealing with the complexity of real auditory scenes in which grouping and categorization cues are not immediately available; however, see (Yildiz and Kiebel 2011). Feedback connections are pervasive within the auditory system, including all stages of the subcortical system, yet to our knowledge no models include such connections. Although fMRI results are useful for identifying regional involvement, detailed understanding of the neural circuitry involved in auditory perceptual organization is sketchy, and the neural representations of auditory objects and perceptual organization are unknown. Even the role of primary auditory cortex remains something of a mystery, e.g., see Nelken et al. (2003) and Griffiths et al. (2004); perhaps studying the switching of perceptual awareness between different representations in awake behaving animals will help to elucidate the representations and processing strategies adopted by cortex.

Cross-References

- ▶ Auditory Event-Related Potentials
- ▶ Stimulus-Specific Adaptation, Models

References

Agus TR, Thorpe SJ, Pressnitzer D (2010) Rapid formation of robust auditory memories: insights from noise. Neuron 66(4):610–618

Alain C, Schuler BM, McDonald KL (2002) Neural activity associated with distinguishing concurrent auditory objects. J Acoust Soc Am 111(2):990–995

Anstis S, Saida S (1985) Adaptation to auditory streaming of frequency-modulated tones. J Exp Psychol Hum Percept Perform 11:257–271

Assmann PF, Summerfield Q (1989) Modeling the perception of concurrent vowels: vowels with the same fundamental frequency. J Acoust Soc Am 85(1):327–338

Assmann PF, Summerfield Q (1990) Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. J Acoust Soc Am 88(2):680–697

Bar M (2007) The proactive brain: using analogies and associations to generate predictions. Trends Cogn Sci 11(7):280–289

Bar-Yosef O, Rotman Y, Nelken I (2002) Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. J Neurosci 22(19):8619–8632

Beauvois MW, Meddis R (1991) A computer model of auditory stream segregation. Q J Exp Psychol A 43(3):517–541

Bee MA (2012) Sound source perception in anuran amphibians. Curr Opin Neurobiol 22(2):301–310

Bee MA, Klump GM (2005) Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences. Brain Behav Evol 66(3):197–214

Bee MA, Micheyl C, Oxenham AJ, Klump GM (2010) Neural adaptation to tone sequences in the songbird forebrain: patterns, determinants, and relation to the build-up of auditory streaming. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 196(8):543–557

Bendixen A, Schröger E, Winkler I (2009) I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. J Neurosci 29(26):8447–8451

Bendixen A, Denham SL, Gyimesi K, Winkler I (2010a) Regular patterns stabilize auditory streams. J Acoust Soc Am 128(6):3658–3666

Bendixen A, Jones SJ, Klump G, Winkler I (2010b) Probability dependence and functional separation of the object-related and mismatch negativity event-related potential components. NeuroImage 50(1):285–290

Bendixen A, Bőhm TM, Szalárdy O, Mill R, Denham SL, Winkler I (2012) Different roles of similarity and predictability in auditory stream segregation. J Learn Percept. (in press)

Bertrand O, Tallon-Baudry C (2000) Oscillatory gamma activity in humans: a possible role for object representation. Int J Psychophysiol 38(3):211–223

Bőhm TM, Shestopalova L, Bendixen A, Andreou AG, Georgiou J, Garreau G, Pouliquen P, Cassidy A, Denham SL, Winkler I (2012) Spatial location of sound sources biases auditory stream segregation but their motion does not. J Learn Percept. (in press)

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT, Cambridge, MA

Brown GJ, Wang DL (eds) (2006) Neural and perceptual modelling. Computational auditory scene analysis: principles, algorithms, and applications. Wiley/IEEE Press, Chichester

Brunswik E (1955) Representative design and probabilistic theory in a functional psychology. Psychol Rev 62(3):193–217

Carlyon RP (2004) How the brain separates sounds. Trends Cogn Sci 8(10):465–471

Ciocca V (2008) The auditory organization of complex sounds. Front Biosci 13:148–169
Culling JF, Summerfield Q (1995) Perceptual separation of concurrent speech sounds: absence of across-frequency grouping by common interaural delay. J Acoust Soc Am 98(2 Pt 1):785–797
Cusack R (2005) The intraparietal sulcus and perceptual organization. J Cogn Neurosci 17(4):641–651
Darwin CJ, Carlyon RP (1995) Auditory grouping. In: Moore BCJ (ed) The handbook of perception and cognition: hearing, vol 6. Academic, London, pp 387–424
Darwin CJ, Hukin RW, al-Khatib BY (1995) Grouping in pitch perception: evidence for sequential constraints. J Acoust Soc Am 98(2 Pt 1):880–885
de Cheveigne A, McAdams S, Laroche J, Rosenberg M (1995) Identification of concurrent harmonic and inharmonic vowels: a test of the theory of harmonic cancellation and enhancement. J Acoust Soc Am 97(6): 3736–3748
Deike S, Heil P, Böckmann-Barthel M, Brechmann A (2012) The build-up of auditory stream segregation: a different perspective. Front Psychol 3:461
Denham SL, Winkler I (2006) The role of predictive models in the formation of auditory streams. J Physiol Paris 100(1–3):154–170
Denham SL, Gymesi K, Stefanics G, Winkler I (2012) Multistability in auditory stream segregation: the role of stimulus features in perceptual organisation. J Learn Percept. (in press)
Duncan J, Humphreys G (1989) Visual search and stimulus similarity. Psychol Rev 96:433–458
Elhilali M, Shamma SA (2008) A cocktail party with a cortical twist: how cortical mechanisms contribute to sound segregation. J Acoust Soc Am 124(6): 3751–3771
Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61(2):317–329
Fay R (2009) Soundscapes and the sense of hearing of fishes. Integr Zool 4(1):26–32
Fishman YI, Arezzo JC, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. J Acoust Soc Am 116(3):1656–1670
Fowler CA, Rosenblum LD (1990) Duplex perception: a comparison of monosyllables and slamming doors. J Exp Psychol Hum Percept Perform 16(4):742–754
Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. Philos Trans R Soc Lond Ser B Biol Sci 364(1521):1211–1221
Griffiths TD, Warren JD (2004) What is an auditory object? Nat Rev Neurosci 5(11):887–892
Griffiths TD, Warren JD, Scott SK, Nelken I, King AJ (2004) Cortical processing of complex sound: a way forward? Trends Neurosci 27(4):181–185
Grimault N, Bacon SP, Micheyl C (2002) Auditory stream segregation on the basis of amplitude-modulation rate. J Acoust Soc Am 111(3):1340–1348
Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25(22):5382–5388
Hartmann WM, Johnson D (1991) Stream segregation and peripheral channeling. Music Percept 9(2):153–183
Haykin S, Chen Z (2005) The cocktail party problem. Neural Comput 17(9):1875–1902
Helmholtz H (1885) On the sensations of tone as a physiological basis for the theory of music. Longmans, Green, London
Hupe JM, Pressnitzer D (2012) The initial phase of auditory and visual scene analysis. Philos Trans R Soc Lond Ser B Biol Sci 367(1591):942–953
Jones MR (1976) Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychol Rev 83:323–355
Jones MR, Kidd G, Wetzel R (1981) Evidence for rhythmic attention. J Exp Psychol Hum Percept Perform 7: 1059–1073
Kashino M, Kondo HM (2012) Functional brain networks underlying perceptual switching: auditory streaming and verbal transformations. Philos Trans R Soc Lond Ser B Biol Sci 367(1591):977–987
Kayser C, Petkov CI, Lippert M, Logothetis NK (2005) Mechanisms for allocating auditory attention: an auditory saliency map. Curr Biol 15(21):1943–1947
Köhler W (1947) Gestalt psychology: an introduction to new concepts in modern psychology. Liveright Publishing Corporation, New York
Kondo HM, Kashino M (2009) Involvement of the thalamocortical loop in the spontaneous switching of percepts in auditory streaming. J Neurosci 29(40): 12695–12701
Kovacs I, Papathomas TV, Yang M, Feher A (1996) When the brain changes its mind: interocular grouping during binocular rivalry. Proc Natl Acad Sci U S A 93(26): 15508–15511
Kubovy M, Van Valkenburg D (2001) Auditory and visual objects. Cognition 80(1–2):97–126
Laing CR, Chow CC (2002) A spiking neuron model for binocular rivalry. J Comput Neurosci 12(1):39–53
Large EW, Jones MR (1999) The dynamics of attending: how people track time-varying events. Psychol Rev 106:119–159
Leopold DA, Logothetis NK (1999) Multistable phenomena: changing views in perception. Trends Cogn Sci 3(7):254–264
Levelt WJM (1968) On binocular rivalry. Mouton, Paris
McCabe SL, Denham MJ (1997) A model of auditory streaming. J Acoust Soc Am 101(3):1611–1621
McDermott JH, Oxenham AJ (2008) Music perception, pitch, and the auditory system. Curr Opin Neurobiol 18(4):452–463
McDermott JH, Wrobleski D, Oxenham AJ (2011) Recovering sound sources from embedded repetition. Proc Natl Acad Sci U S A 108(3):1188–1193
Micheyl C, Tian B, Carlyon RP, Rauschecker JP (2005) Perceptual organization of tone sequences in

the auditory cortex of awake macaques. Neuron 48(1): 139–148
Micheyl C, Carlyon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Courtenay Wilson E (2007) The role of auditory cortex in the formation of auditory streams. Hear Res 229(1–2):116–131
Mill R, Bőhm T, Bendixen A, Winkler I, Denham SL (2013) Competition and cooperation between fragmentary event predictors in a model of auditory scene analysis. PLoS Comput Biol. (in press)
Miller GA, Licklider JCR (1950) The intelligibility of interrupted speech. J Acoust Soc Am 22:167–173
Moore BCJ, Gockel HE (2002) Factors influencing sequential stream segregation. Acta Acust 88:320–333
Moore BC, Gockel HE (2012) Properties of auditory stream formation. Philos Trans R Soc Lond Ser B Biol Sci 367(1591):919–931
Moore BC, Glasberg BR, Peters RW (1986) Thresholds for hearing mistuned partials as separate tones in harmonic complexes. J Acoust Soc Am 80(2):479–483
Näätänen R, Winkler I (1999) The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull 125(6):826–859
Näätänen R, Gaillard AWK, Mäntysalo S (1978) Early selective attention effect on evoked potential reinterpreted. Acta Psychol 42:313–329
Nager W, Teder-Sälejärvi W, Kunze S, Münte TF (2003) Preattentive evaluation of multiple perceptual streams in human audition. Neuroreport 14(6):871–874
Nakajima Y, Sasaki T, Kanafuka K, Miyamoto A, Remijn G, ten Hoopen G (2000) Illusory recouplings of onsets and terminations of glide tone components. Percept Psychophys 62(7):1413–1425
Nelken I (2008) Processing of complex sounds in the auditory system. Curr Opin Neurobiol 18(4):413–417
Nelken I, Fishbach A, Las L, Ulanovsky N, Farkas D (2003) Primary auditory cortex of cats: feature detection or something else? Biol Cybern 89(5):397–406
Oertel D, Fay RR, Popper AN (2002) Integrative functions in the mammalian auditory pathway. Springer, New York
Pearl J (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, San Mateo
Pressnitzer D, Hupe JM (2006) Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. Curr Biol 16(13): 1351–1357
Rand TC (1974) Letter: dichotic release from masking for speech. J Acoust Soc Am 55(3):678–680
Rensink RA (2000) Seeing, sensing, and scrutinizing. Vis Res 40(10–12):1469–1487
Riecke L, Van Opstal AJ, Formisano E (2008) The auditory continuity illusion: a parametric investigation and filter model. Percept Psychophys 70(1):1–12
Roberts B, Glasberg BR, Moore BC (2002) Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. J Acoust Soc Am 112(5 Pt 1):2074–2085
Schadwinkel S, Gutschalk A (2011) Transient bold activity locked to perceptual reversals of auditory streaming in human auditory cortex and inferior colliculus. J Neurophysiol 105(5):1977–1983
Schwartz JL, Grimault N, Hupe JM, Moore BC, Pressnitzer D (2012) Multistability in perception: binding sensory modalities, an overview. Philos Trans R Soc Lond Ser B Biol Sci 367(1591):896–905
Shamma SA, Elhilali M (2013)
Shamma SA, Micheyl C (2010) Behind the scenes of auditory perception. Curr Opin Neurobiol 20(3):361–366
Shamma SA, Elhilali M, Micheyl C (2011) Temporal coherence and attention in auditory scene analysis. Trends Neurosci 34(3):114–123
Shamma S, Elhilali M, Ma L, Micheyl C, Oxenham AJ, Pressnitzer D, Yin P, Xu Y (2013) Temporal coherence and the streaming of complex sounds. Adv Exp Med Biol 787:535–543
Shinn-Cunningham BG (2008) Object-based auditory and visual attention. Trends Cogn Sci 12(5):182–186
Shpiro A, Moreno-Bote R, Rubin N, Rinzel J (2009) Balance between noise and adaptation in competition models of perceptual bistability. J Comput Neurosci 27(1):37–54
Snyder JS, Alain C (2007) Toward a neurophysiological theory of auditory stream segregation. Psychol Bull 133(5):780–799
Summerfield Q, Assmann PF (1991) Perception of concurrent vowels: effects of harmonic misalignment and pitch-period asynchrony. J Acoust Soc Am 89(3): 1364–1377
Sussman ES, Ritter W, Vaughan HG Jr (1999) An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology 36(1):22–34
Szalárdy O, Bendixen A, Tóth D, Denham SL, Winkler I (2012) Modulation-frequency acts as a primary cue for auditory stream segregation. J Learn Percept. (in press)
Szalárdy O, Bőhm T, Bendixen A, Winkler I (2013) Perceptual organization affects the processing of incoming sounds: an ERP study. Biol Psychol. (in press)
Takegata R, Brattico E, Tervaniemi M, Varyagina O, Naatanen R, Winkler I (2005) Preattentive representation of feature conjunctions for concurrent spatially distributed auditory objects. Brain Res Cogn Brain Res 25(1):169–179
Teki S, Chait M, Kumar S, von Kriegstein K, Griffiths TD (2011) Brain bases for auditory stimulus-driven figure-ground segregation. J Neurosci 31(1):164–171
Treisman A (1998) Feature binding, attention and object perception. Philos Trans R Soc Lond Ser B Biol Sci 353:1295–1306
van Ee R (2009) Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: evidence from serial correlations in perceptual bistability. J Opt Soc Am A Opt Image Sci Vis 26(12):2612–2622
van Noorden LPAS (1975) Temporal coherence in the perception of tone sequences. Doctoral dissertation, Technical University Eindhoven

Vliegen J, Oxenham AJ (1999) Sequential stream segregation in the absence of spectral cues. J Acoust Soc Am 105(1):339–346
von Ehrenfels C (1890) Über Gestaltqualitäten (English "on the qualities of form"). Vierteljahrsschr Wiss Philos 14:249–292
von Kriegstein K, Smith DR, Patterson RD, Ives DT, Griffiths TD (2007) Neural representation of auditory size in the human voice and in sounds from other resonant sources. Curr Biol 17(13):1123–1128
Wang DL, Brown GJ (2006) Computational auditory scene analysis: principles, algorithms, and applications. Wiley/IEEE Press, New York
Wang DL, Chang PS (2008) An oscillatory correlation model of auditory streaming. Cogn Neurodyn 2:7–19
Warren RM, Wrightson JM, Puretz J (1988) Illusory continuity of tonal and infratonal periodic sounds. J Acoust Soc Am 84(4):1338–1342
Weiss Y, Simoncelli EP, Adelson EH (2002) Motion illusions as optimal percepts. Nat Neurosci 5(6):598–604
Wertheimer M (1912) Experimentelle Studien über das Sehen von Bewegung. Z Psychol 60
Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ (2007) Cortical FMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming. J Neurophysiol 97(3):2230–2238
Winkler I (2007) Interpreting the mismatch negativity. J Psychophysiol 21:147–163
Winkler I (2010) In search for auditory object representations. In: Winkle I, Czigler I (eds) Unconscious memory representations in perception: processes and mechanisms in the brain. John Benjamins, Amsterdam, pp 71–106
Winkler I, Cowan N (2005) From sensory to long-term memory: evidence from auditory memory reactivation studies. Exp Psychol 52(1):3–20
Winkler I, Czigler I (2012) Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. Int J Psychophysiol 83(2):132–143
Winkler I, Sussman E, Tervaniemi M, Horváth J, Ritter W, Näätänen R (2003a) Preattentive auditory context effects. Cogn Affect Behav Neurosci 3(1):57–77
Winkler I, Teder-Salejarvi WA, Horváth J, Näätänen R, Sussman E (2003b) Human auditory cortex tracks task-irrelevant sound sources. Neuroreport 14(16): 2053–2056
Winkler I, Czigler I, Sussman E, Horvath J, Balazs L (2005a) Preattentive binding of auditory and visual stimulus features. J Cogn Neurosci 17(2):320–339
Winkler I, Takegata R, Sussman E (2005b) Event-related brain potentials reveal multiple stages in the perceptual organization of sound. Brain Res Cogn Brain Res 25(1):291–299
Winkler I, van Zuijen TL, Sussman E, Horvath J, Naatanen R (2006) Object representation in the human auditory system. Eur J Neurosci 24(2):625–634
Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn Sci 13(12): 532–540
Winkler I, Denham S, Mill R, Bohm TM, Bendixen A (2012) Multistability in auditory stream segregation: a predictive coding view. Philos Trans R Soc Lond Ser B Biol Sci 367(1591):1001–1012
Yildiz IB, Kiebel SJ (2011) A hierarchical neuronal model for generation and online recognition of birdsongs. PLoS Comput Biol 7(12):e1002303
Zhuo G, Yu X (2011) Auditory feature binding and its hierarchical computational model. In: Third international conference on artificial intelligence and computational intelligence. Springer
Zwicker E, Fastl H (1999) Psychoacoustics. Facts and models. Springer, Heidelberg/New York

Auditory Precedence Effect

Barbara Shinn-Cunningham
Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA

Synonyms

Law of the first wavefront

Definition

The precedence effect is a well-studied phenomenon in spatial hearing that is related to how we localize sounds accurately in everyday settings. Specifically, when two sound sources reach a listener close together in time, listeners often hear a single "fused" image whose perceived direction is near the location of the first-arriving sound.

Detailed Description

The Effects of Room Acoustics on Auditory Spatial Cues

The signals reaching the listener's ears directly from a sound source convey information about the source's location (Blauert 1997; Schnupp et al. 2010). However, in ordinary

settings, soon after the direct sound reaches the listener, reflected sound arrives from random directions, coming off of walls, floors, and other reflective surfaces. This reflected sound energy adds acoustically to the direct sound before entering each ear, changing the total signal reaching the ear (e.g., see Allen and Berkley 1979). The content of this reflected sound energy depends on the geometry of the listening space, the positions of objects in the space, and the location and orientation of the listener. Moreover, the reflected sound energy differs at the left and right ears. As a result, this late-arriving sound distorts the spatial information conveyed by the direct sound, degrading the spatial cues in the total signal. As a result, the acoustic location information in the signals a listener hears in everyday settings is less reliable after an initial "clean glimpse" of the direct sound that happens at sound onset, before the reflected energy arrives.

The Precedence Effect

Perceptually, judgments of the direction of a sound source depend strongly on spatial information in the onset of sound and relatively weakly on spatial information in later-arriving portions of sound (e.g., see Brown and Stecker 2010). This phenomenon is known as the precedence effect (Wallach et al. 1949; Zurek 1987; Litovsky et al. 1999). Given that in ordinary listening spaces the onset of the sound has the most reliable information about source direction, the precedence effect is thought to be one of the reasons why listeners are relatively good at judging source location even when listening in rooms.

The precedence effect has been studied extensively with pairs of clicks (one leading and one lagging); for such brief stimuli, the precedence effect is strongest when the leading click precedes the lagging click by 1–5 ms and then rapidly becomes weaker (so that listeners start to hear the second click as a separate event and then begin to localize it with increasing accuracy; see Blauert 1997). For more "natural" sounds, like speech or music, the precedence effect persists for tens of ms. Many researchers argue that the precedence effect has a longer time course for ongoing signals because they can be thought of as containing multiple "onsets" due to local energy fluctuations, each of which can add to the precedence effect (e.g., see Zurek 1980).

Mechanisms of the Precedence Effect

Although the precedence effect is often discussed as a single psychophysical phenomenon, many different mechanisms likely contribute to the dominance of early spatial information on later-arriving information. For instance, the most peripheral portion of the auditory system, the auditory nerve, responds more vigorously at the onset of a sound than to later-arriving portions of a sound. This peripheral adaptation helps explain why a lagging sound that arrives within a few milliseconds of a leading sound does not convey strong spatial cues (Hartung and Trahiotis 2001). However, there are numerous studies that show that the perceptual dominance of the leading sound extends beyond very brief lead-lag delays that can be fully explained by peripheral adaptation. It is likely that microcircuitry in the brainstem contributes to the precedence effect at longer lead-lag delays through some type of inhibition triggered by the leading sound (Xia and Shinn-Cunningham 2011).

Cross-References

- ▶ Sound Localization and Experience-Dependent Plasticity
- ▶ Sound Localization in Mammals and Models

References

Allen JB, Berkley DA (1979) Image method for efficiently simulating small-room acoustics. J Acoust Soc Am 65:943–950

Blauert J (1997) Spatial hearing (2e). MIT Press, Cambridge, MA, pp 201–287

Brown AD, Stecker GC (2010) Temporal weighting of interaural time and level differences in high-rate click trains. J Acoust Soc Am 128:283–292

Hartung K, Trahiotis C (2001) Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. J Acoust Soc Am 110:1505–1513

Litovsky RY, Colburn HS, Yost WA, Guzman SJ (1999) The precedence effect. J Acoust Soc Am 106:1633–1654
Schnupp J, Nelken I, King A (2010) Auditory neuroscience. MIT Press, Cambridge, MA, pp 177–222
Wallach H, Newman EB, Rosenzweig MR (1949) The precedence effect in sound localization. Am J Psychol 52:315–336
Xia J, Shinn-Cunningham BG (2011) Isolating mechanisms that influence measures of the precedence effect: theoretical predictions and behavioral tests. J Acoust Soc Am 130:866–882
Zurek PM (1980) The precedence effect and its possible role in the avoidance of interaural ambiguities. J Acoust Soc Am 67:952–964
Zurek PM (1987) The precedence effect. In: Yost WA, Gourevitch GA (eds) Directional hearing. Springer, New York, pp 85–105

Auditory Processing in Insects

R. Matthias Hennig and Bernhard Ronacher
Department of Biology, Humboldt-Universität zu Berlin, Berlin, Germany

Synonyms

Auditory pathway; Hearing

Definition

Auditory processing in insects serves to extract relevant information from acoustic signals about identity and location of acoustic objects, usually in the context of mate attraction and predator avoidance. For this goal, insects process spectral information from the carrier frequency of a signal and obtain temporal information from the sound's amplitude modulation pattern (the envelope). Auditory processing in insects is constrained by size in several aspects: first, the signal often contains ultrasonic frequencies due to small sender size; second, for localization, insects have only poor directional cues because of the small distance between their ears; and third, due to their small brains, the auditory processing capacity is limited to a small number of neurons.

Detailed Description

Overview and Background

Large Diversity of Hearing Insects and their Ears

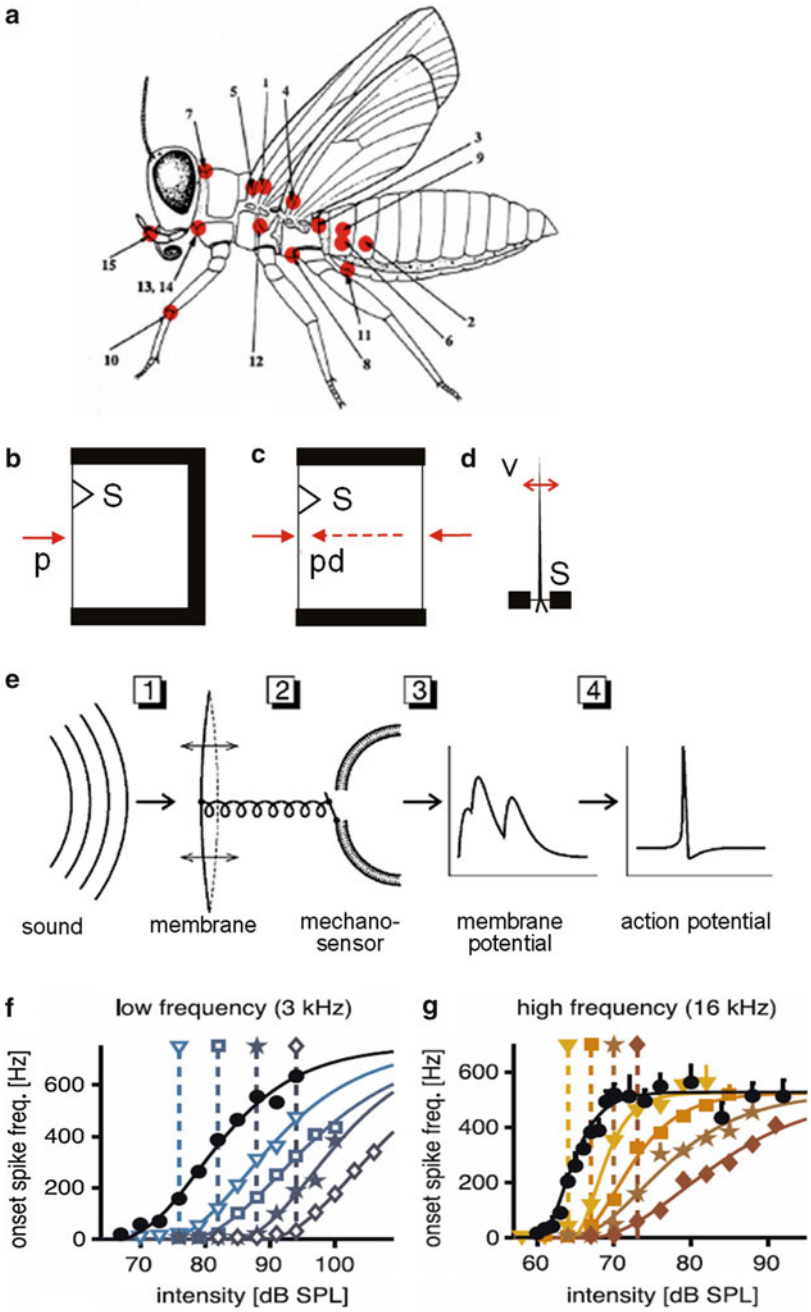
The sense of hearing has evolved in vertebrates and arthropods, and hearing organs are known from several orders of insects. Ears have evolved in rather different locations of their body, not only on the head but also on the thorax, abdomen, and even wings and legs (Fig. 1a, Fullard and Yack 1993). Consequently, numerous neuronal substrates for the processing of acoustic information are known that evolved from a mechanosensory modality employed in a different context (proprioception by chordotonal organs: Meier and Reichert 1990; van Staaden and Römer 1998). In insects, two types of ears are known that are sensitive to two different physical attributes of sound, that is, the movement of particles in an elastic medium: tympanal ears that are specialized to sense changes in sound pressure and antennae or filiform hairs that are sensitive to particle velocity (Fig. 1b–d, Michelsen 1979).

Goals of Hearing and Auditory Processing

The ability to hear sound and to extract relevant information from an acoustic signal has evolved in three functional contexts. Notably, the production of sound alone is not a sufficient indicator of hearing ability as many insects produce defensive sounds when threatened (e.g., Bura et al. 2011). Conversely, numerous insects can hear sound but are themselves mute (Riede 1987; Riede et al. 1990).

Acoustic Communication

In many species the perception of acoustic signals occurs in the context of acoustic communication for mate recognition and mate localization. The most prominent examples stem from grasshoppers, crickets, and bushcrickets (orthoptera), but cicadas and moths have also evolved elaborate acoustic signals. A major goal of intersexual acoustic communication is to discriminate conspecific from heterospecific signals which helps to avoid fitness losses (Fig. 2). Sexual selection by



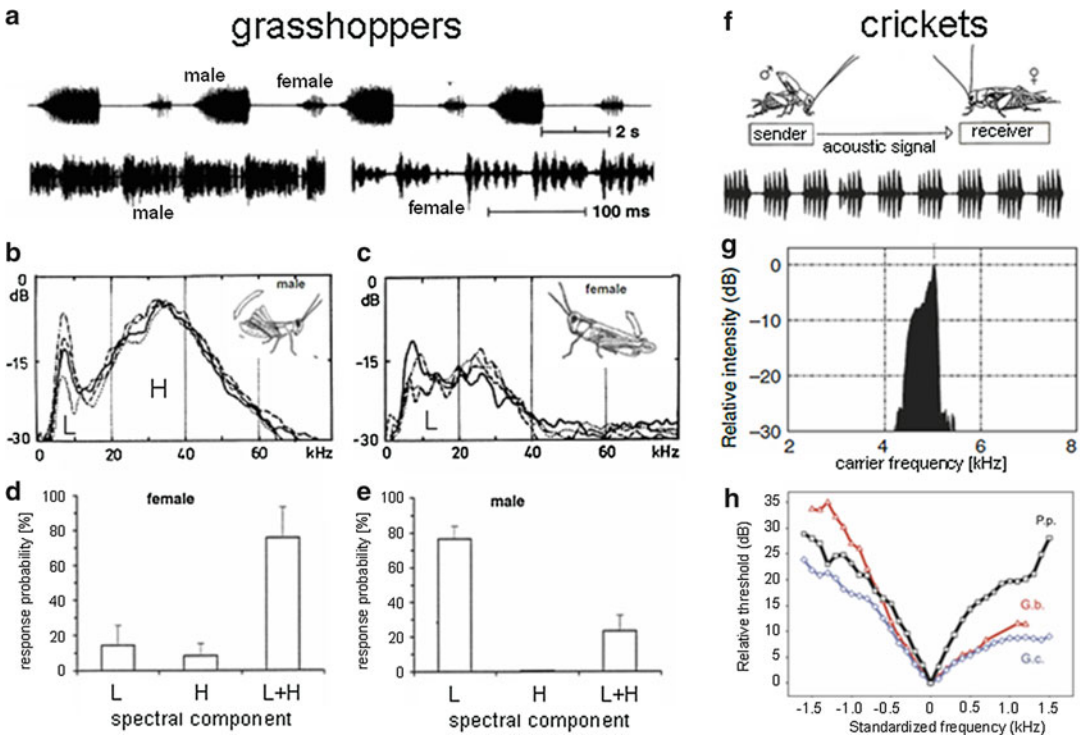
Auditory Processing in Insects, Fig. 1 Insect ears and sound transduction. (**a**) Evolution of insect ears in different body locations (**b**–**d**). Types of ears: (**b**) sound pressure receiver (mammals and humans); (**c**) pressure difference receiver (insects); (**d**) sound velocity receiver (filiform hair of insects); *S* sensory cells, *p* sound pressure, *pd* sound pressure difference, and *v* particle velocity. (**e**) Sensory transduction at a tympanic ear: 1, sound wave; 2, vibration of membrane and movement of mechanosensitive ion channels due to sound pressure; and 3–4, membrane potential and elicited action potential in sensory cell. (**f**, **g**) Intensity-response curves of an auditory neuron in crickets for different background intensities and different carrier frequencies (**f**, 3 kHz; **g**, 16 kHz). At increasing background intensities as indicated by the *top symbols* in (**f**, **g**), the response curves shift to higher intensities. ((**a**) Modified from Fullard and Yack 1993, with permission; see also there for explanation of numbers. (**b**–**d**) From Penzlin 2005, with permission. (**e**) From Gollisch and Herz 2005, with permission. (**f**, **g**) From Hildebrandt et al. 2011, with permission)

female choice for song signals of particularly attractive mating partners is also well known (von Helversen and von Helversen 1994; Andersson and Simmons 2006). Insects employ very stereotyped signals for acoustic communication, the production and recognition of which has a strict innate basis (Bentley and Hoy 1972, von Helversen and von Helversen 1975a, b, 1987). For this reason, the acoustic communication of grasshoppers, crickets, and bushcrickets has served as a model system to study the mechanisms of neuronal processing within the auditory pathway by carefully designed behavioral experiments and recordings of neuronal activity of

single cells (Schildberger and Elsner 1994; Huber et al. 1989; Gerhardt and Huber 2002).

Since the nervous system of insects is small, all computations must be performed by relatively few neurons. This likely imposes a strong pressure for an efficient and sparse representation of the external world. A big advantage of investigating nervous system processes via the study of communication signals is that an animal's "external world" can be reduced to a manageable set of relevant stimuli, the stimulus space. This rather small set of highly relevant stimuli allows one to characterize how the stimuli are represented within the nervous system and how these representations are transformed at different stages of processing.

To allow for successful communication, signal properties must be matched to the sensory characteristics of receivers. A straightforward example is the frequency tuning of a cricket female's ears to the narrowband signals of the males (see Fig. 2g, h). However, since in many species the temporal pattern of communication signals conveys the species-specific information, sender and receiver should be matched not only for carrier frequency but also for temporal characteristics.



Auditory Processing in Insects, Fig. 2 Acoustic communication in grasshoppers and crickets. (**a**, **f**) A *sender*, the male, produces a specific sound *signal* that is *transmitted* and perceived by a *receiver*, the female. (**a**) In some grasshoppers, the female responds to the male with her own song (*upper trace*). Songs of males and females have different pulse shapes (*lower traces*). (**b**, **c**) The power spectra differ between males (**b**) and females (**c**) in relative content of low (L)- and high (H)- frequency components. (**d**, **e**) Females and males respond selectively to the relative content of low- and high-frequency components in a signal. (**d**) Females prefer the combination of low and high components as they occur in the male song (as in **b**). (**e**) Males prefer signals with only low-frequency components as they are typical for a female song (as in **c**). (**g**) The spectral content of a song of a cricket as in (**f**) is narrowly tuned to a peak frequency (5.0 kHz). (**h**) Auditory tuning for the specific carrier frequency of a conspecific song for different species of crickets (frequency axis was normalized to the respective best frequency; *P.p. Paroecanthus podagrosus*, *G.b. Gryllus bimaculatus*, *G.c. G. campestris*). Note the differences in the sharpness of tuning ((**a**–**e**) From von Helversen and von Helversen 1997; (**f**) modified from Huber 1992; (**g**) modified from Montealegre-Z et al. 2011; and (**h**) from Schmidt et al. 2011, with permission)

Here, a new problem arises: temperature. Since the function of neurons and muscles is strongly temperature-dependent (Janssen 1992; Robertson and Money 2012) and in general insects' body temperatures vary directly with ambient temperature, a signal's temporal pattern will vary with ambient temperature. This creates a recognition problem for the receiver's auditory system, in particular, if sender and receiver differ in their body temperatures (von Helversen and von Helversen 1987; Gerhardt and Huber 2002; Ronacher et al. 2004).

Detection and Avoidance of Predators

Although hearing in insects evolved more than 200 million years ago (MYA), the appearance of bats about 60 MYA sparked an explosion of insect species with ears (Fig. 1a, Fullard and Yack 1993; Hoy et al. 1998; Stumpner and von Helversen 2001). While some species modified already-existing ears for the detection of ultrasound, numerous other insects that had previously lacked hearing evolved ultrasonic ears at this time (Miller and Surlykke 2001). For moths this acquired sensory ability may even have served as a starting point for the evolution of intraspecific acoustic communication. Within their auditory pathways, insects employ different neuronal substrates for the processing of intraspecific signals and sounds from predators (see below: categorical perception in crickets and parallel processing of information). Some species perform stream segregation within individual neurons by spectral and temporal cues (Schul and Sheridan 2006). Numerous species of grasshoppers have ears and hearing abilities, but produce no sound during mate attraction. For these species, hearing likely serves for predator detection only, and predation was the selection pressure under which ears evolved or were conserved (Riede 1987; Riede et al. 1990; Lehmann et al. 2007).

Host Finding

Hearing has evolved independently at least twice in parasitic flies in the context of finding hosts for their eggs. The performance of flies in localizing their host by its sound signal is most impressive, in view of their tiny ears located underneath the head (Robert et al. 1996; Lakes-Harlan et al. 1999). With these tympanal ears flies can detect and localize not only the broadband sounds of bushcrickets and cicadas but also the pure tone signals emitted by crickets. For that goal flies possess sharply tuned hearing, as is evident from the sensitivity of their sensory receptors and auditory interneurons (Stumpner and Lakes-Harlan 1996; Robert and Hoy 1998; Robert and Göpfert 2002). Another example of host finding via auditory cues is a bloodsucking corethrellid fly that is attracted by frog calls (Bernal et al. 2006).

Two Types of Ears in Insects and their Constraints by Size and Signals

Generally, two principal types of ears are known: particle velocity receivers and sound pressure receivers (Fig. 1b–d, Michelsen 1979; Faure et al. 2009). (1) Particle velocity receivers exploit the vector component of sound particles close to the sound source. The antennae (arista) of flies, especially of *Drosophila* and mosquitoes, are a well-investigated model system for sound perception, auditory transduction, and active sensing (Robert and Göpfert 2002). Many insects, as well as arthropods in general, employ filiform hairs (located on the abdomen, cerci, legs, or other parts of the body, Fig. 1d) of different length to detect sounds of predators and prey (Barth 2002). A general property of particle velocity receivers is their limitation to lower frequency ranges (<500 Hz). Low frequency signals from small senders commonly have low amplitude and thus small range (e.g., wing movements by *Drosophila*), and the perception of particle velocity is then usually limited to the near field of the sender (a distance of a few wavelengths). Since the vector component of sound is perceived, these ears also provide directional information (Michelsen 1979; Faure et al. 2009). Their sensitivity matches or even surpasses that of tympanal ears (Robert and Göpfert 2002). (2) The perception of sound pressure is mediated by tympanal ears and follows the same principles as in vertebrates including humans (Fig. 1b, c, Montealegre et al. 2012). Tympanal ears in insects arose from the cuticular surface of their exoskeleton under which large air sacs derived from tracheal tubes, the respiratory

system of insects, were located. Since insects possess an abundance of mechanosensory proprioreceptors for monitoring the strain and movement of their cuticle, auditory organs were prone to evolve from chordotonal organs in almost any part of the body (e.g., the legs, thorax, abdomen, and wings, Fig. 1a, Fullard and Yack 1993; van Staaden and Römer 1998). Tympana are usually small, and the sensitivity of the ears often extends into the ultrasonic frequency range. Sound localization via tympanal ears in larger vertebrates (Fig. 1b) requires the computation of interaural intensity and time differences as pressure receivers do not respond to the vector component of sound (Brown 1994; Yost 2000). Due to the small size of insects, intensity differences and in particular interaural time differences are too small to be exploited directly. Insects circumvent this problem by using tympanic *pressure difference* receivers, in which an internal connection by tracheal tubes exists between both ears (Fig. 1c). Small frogs, lizards, and birds face similar problems and also have developed pressure difference receivers. In these ears, the vibration amplitude of the tympanum is determined not only by the sound pressure at the outer side but also by the sound traveling through the body to the inside of the tympanum (Autrum 1942). The resulting pressure difference between the inside and outside will then determine the tympanal vibration. The amplitude of these vibrations depends on sound direction, because the phase angle of a given sound frequency at both sides of the tympanum is also dependent on the direction of the incident sound wave (Michelsen et al. 1994).

Themes of Auditory Processing in Insects

A major challenge in summarizing the capacities for auditory processing of insects results from the overwhelming diversity of hearing species, of functional ears, and of the different designs of auditory pathways. The subchapters below therefore give only a brief overview of the computational capabilities of insects for different tasks and under different constraints.

For auditory processing, insects exploit numerous general principles of sensory processing that are well known from other modalities and from vertebrates, including mammals. Among these are the capacity for sound frequency analysis by a traveling wave, tonotopic representations and formation of internal neuronal maps, parallel processing of information, the timing and balance of excitation and inhibition for feature extraction, lateral and contralateral inhibition for contrast enhancement, transformation of coding from a temporal code to a place code, burst coding, resonant properties of neurons, selective attention, and even stream segregation.

However, insects face constraints on the computational power provided by their small brains. The concept of identified neurons, in which individual neurons could be identified by their morphology and physiology, arose from neurobiological research in insects and other arthropods (Huber and Markl 1983). Computations performed by thousands of neurons in mammals may find their counterpart in a single identifiable neuron of an insect, which illustrates an impressive compression of function (e.g., contralateral inhibition for directional hearing is mediated by the lateral superior olive in mammals and by a single local interneuron, ON1, in crickets and bushcrickets (Grothe 2000; Selverston et al. 1985; Römer and Krusch 2000).

Notably, the processing and coding capacity of insect nervous systems is restricted to relevant tasks. Peripheral filters and computations for instance aid in reducing the required processing power. Therefore, the ears of insects and their auditory pathways are by no means all-purpose devices, and processing is rather specific to function. Examples include crickets that distinguish only 2 categories of sound (mate signals and predators, Hoy 1989) and the tympanic ear of a moth that is equipped with only 2 sensory cells for bat detection (Boyan and Fullard 1988). Generally, insects also employ simple algorithms for processing at the cost of acuity, for example, by computing acoustic hemispheres rather than localizing the angle of a sound source (von Helversen 1997; Römer and Krusch 2000). Nevertheless, insects are by no means imprecise. At least in some species their performance for temporal resolution in a gap detection task has a precision in the millisecond range, rivaling that

of humans (von Helversen 1972; Prinz and Ronacher 2002).

Basic Steps of Auditory Processing: Transduction and Information Coding by Sensory Neurons

In both tympanal ears and particle velocity receivers, sound induces the vibration of a structure (a thin membrane or a lever on a flexible pivot, Michelsen 1979; Robert and Göpfert 2002). The mechanical oscillation distorts the dendrite of a scolopidial cell, which is attached to the lever, the tympanum, or a tracheal tube (Fig. 1e1, e2). This distortion opens mechanosensitive ion channels, which are not yet fully characterized, and the resulting current transduces the movement into a change of the cell's membrane potential (Fig. 1e3; Gollisch and Herz 2005). Evidently, the membrane potential of sensory neurons cannot follow the fast vibrations of the tympanal membrane in the kilo-Hertz range; rather, the membrane potential depends on the instantaneous sound pressure level. Similar to vertebrates, frequency discrimination occurs according to a frequency-place transformation (Montealegre et al. 2012). Either in the sensory neuron itself or in a downstream neuron, the membrane depolarization is then translated into a series of action potentials, a spike train, which encodes the sound envelope by modulations of the spike rate (Fig. 1e4). Spike rates of auditory afferents can be high, up to 400 Hz or more (Römer 1976).

Although insect and vertebrate ears obviously evolved independently, and the insect mechanoreceptors (scolopidia) differ from auditory hair cells of vertebrates, in the last decade many unexpected commonalities were detected. Recently, active processes were found in insect ears that serve to attain and adjust their formidable sensitivity, similar to the function of outer hair cells in the mammalian cochlea (Robert and Göpfert 2002; Nadrowski et al. 2011). In both locusts and bushcrickets, traveling waves were observed to play an essential role in frequency discrimination. As in the mammalian cochlea, traveling waves exhibit peaks at different locations, depending on sound frequency (Windmill et al. 2005; Hummel et al. 2011; Montealegre et al. 2012). In addition, in central projections of auditory afferents, there is a tonotopic representation of frequencies (Römer 1983; Römer et al. 1988; Stumpner 1996; Stölting and Stumpner 1998). Finally, the development of the auditory organ of *Drosophila* depends on similar genes, e.g., of the atonal family, as the vertebrate ear (Senthilan et al. 2012).

Although the *temporal* resolution of insect ears can match that of vertebrates in some respects (see below), the *frequency resolution* of insects is generally inferior to that of vertebrates. In crickets we find a kind of "categorical response," by which the frequency scale is segmented into two parts: sound pulses with carrier frequencies above 15–20 kHz (up to 100 kHz) evoke an avoidance response, whereas sound pulses with frequencies between 3 and 10 kHz are attractive and evoke a positive steering response when flying (Moiseff et al. 1978; Hoy 1989). Remarkably, however, in an interneuron of a bushcricket, a sharpening of the broader frequency tuning of auditory receptors by frequency-dependent "lateral" inhibition similar to vertebrates has been observed (Stumpner 1997). The sharpness of frequency tuning of auditory neurons is commonly described by the Q_{10dB} score. This gives the ratio between the characteristic frequency (i.e., the frequency of the neuron's lowest threshold) and the width of the tuning curve 10 dB above the lowest threshold. Typical values for insects are between 0.5 and 2.5 and only rarely extend to 3.5 (Hennig et al. 2004), whereas in vertebrates, we find much higher values, between 1 and 25 (and up to 400 in the acoustic foveae of bats, Suga et al. 1997).

The dependency of spike rate on sound intensity is commonly depicted by the f-I curve (firing rate vs. intensity, also rate-intensity curve; Fig. 1f, g). The dynamic range between threshold and saturation indicates the region of discriminable sound intensities. The steepness of the f-I curve determines how well small sound pressure differences can be discriminated – which is important for directional hearing and when fine modulation details of a signal have to be assessed (compare

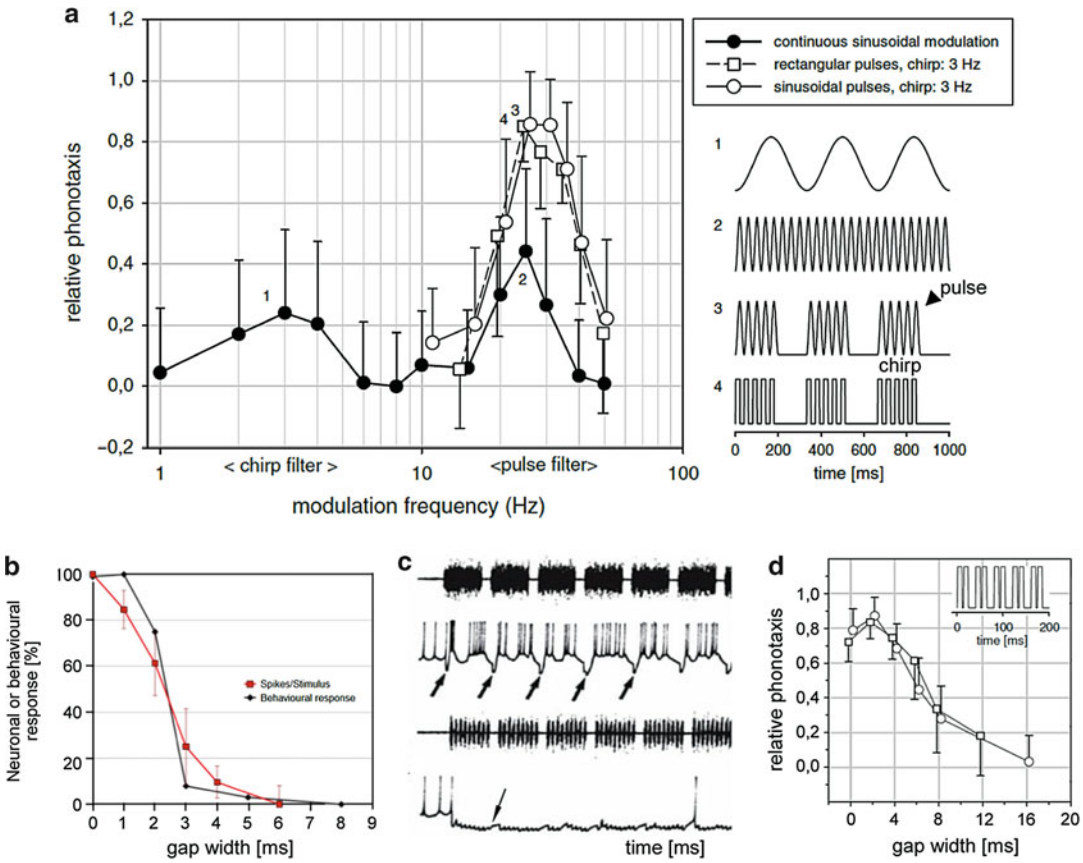
Fig. 1f, g). A steep f-I curve, however, has the disadvantage that it covers only a small part of the relevant sound intensity range which may extend to 100–120 dB SPL (Fig. 1f). (Note that the dB SPL scale is a logarithmic scale that expresses sound pressure relative to 20 μPa; a 100 dB sound has a 10^5 larger sound pressure compared to the human hearing threshold at 1 kHz.) As in other sensory systems, the intensity range problem may be solved by adaptation, by which the f-I curve can be adjusted to the average ambient sound pressure levels (Benda and Herz 2003; Benda and Hennig 2008). Spike frequency adaptation as early as the level of sensory neurons helps to attain a certain degree of intensity invariance that is important for object identification and behavioral decisions (Benda and Hennig 2008). However, in different neuron types we may find different biophysical realizations, ranging from cell-intrinsic spike-triggered adaptation currents to inhibitory inputs and presynaptic adaptation mechanisms (Gollisch and Herz 2004; Hildebrandt et al. 2009). A complementary way to cope with the large range of encountered sound intensities may be a kind of "range fractionation." In locusts and bushcrickets, for example, we find auditory receptors with a rather limited dynamic range of 20–30 dB. Since different receptors exhibit a similarly broad frequency tuning but very different thresholds, intensity discrimination is possible over a broad range (Römer 1976; Römer et al. 1998).

Obviously, the spike trains of sensory neurons are the only information any central nervous system has about events in the external world. In other words, the brain has to infer the structure of the outer world from the spike trains arriving from various sense organs. The very successful stimulus reconstruction methods seek to understand signal processing from the viewpoint of the central nervous system and to infer information about an external stimulus from the spike trains of sensory neurons (Rieke et al. 1997). The basic idea is to reconstruct the stimulus envelope from spike trains that have been recorded in response to that stimulus. By this procedure we can estimate how much information the CNS can obtain about a sensory stimulus and what aspects of the stimulus are lost. For example, from the spike trains of the locusts' auditory afferents, stimuli with large modulation amplitudes can be reconstructed more accurately than stimuli with small modulation depths (Machens et al. 2001). Remarkably, grasshopper songs seem to be matched to this feature of the receiver's auditory periphery. This investigation has further shown that in the very periphery of the auditory pathway, a single sensory neuron transmits a high amount of information, up to 180 bits/s (Machens et al. 2001).

Processing of Signal Envelopes within the Auditory Pathway

Temporal Resolution and Temporal Integration

The communication signals of many species contain fast amplitude modulations that are evaluated by females to assess the attractiveness of potential mates. Hence, we must ask whether there are neuronal constraints that determine the behavioral *limits of temporal resolution*. Two widely used paradigms to investigate these limits are modulation transfer functions (MTF) and gap detection (for reviews see Green 1985; de Boer 1985; Michelsen 1985; Viemeister and Plack 1993; Joris et al. 2004). In the MTF paradigm either random modulations in a certain frequency band or sine wave modulations are used. The latter paradigm can be applied in neurophysiological as well as in behavioral experiments. Stimuli with appropriate carrier frequencies are presented that exhibit sinusoidal amplitude modulations of different frequencies and reveal what range of modulation frequencies the system is able to represent, and if there are specific modulation frequencies to which a system responds particularly well, see Fig. 3a for an example. The black dots in this diagram show the attraction of cricket females to a 4.5 kHz tone that was amplitude modulated at frequencies between 1 and 50 Hz; two regions of enhanced attractiveness around 3 and 30 Hz are obvious (Wendler 1989; Hennig 2009). In spike train recordings one can determine the average spike count



Auditory Processing in Insects, Fig. 3 Temporal resolution of amplitude modulations in grasshoppers and crickets. (**a**) Behavioral modulation transfer function of crickets (filled symbols, stimuli as in 1 and 2 at right). Best responses are obtained if lower and higher modulation frequencies for pulse and chirp (i.e., groups of pulses) are combined in one stimulus as in patterns 3 and 4 at right (open symbols). (**b**) Gap detection in a grasshopper measured in behavioral experiments (*black curve*) and neuronal response (*red curve*, AN4). (**c**) Response of the AN4 neuron to uninterrupted and gap containing sound syllables. This neuron responds to sound onset first with a deep inhibition, an IPSP (*arrows*), followed by excitation and spikes (*upper traces*). In the interrupted stimuli, each onset after a gap triggers the IPSP anew which leads to an effective suppression of spiking (*lower traces*, adapted from Ronacher & Stumpner 1988). (**d**) Gap detection in crickets (behavioral data) ((**a**) From Hennig 2009; (**b**) modified from von Helversen 1972 and Franz and Ronacher 2002; (**c**) modified from Ronacher and Stumpner 1988; and (**d**) from Schneider and Hennig 2012, with permission)

(rate, r-MTF) or evaluate how well the spikes are locked to a period of the stimulus envelope (temporal, t-MTF). r-MTF reveals a neuron's filter properties, e.g., high-pass, low-pass, band-pass, or band-reject features for sound pulse rates. The t-MTF, in contrast, indicates how well fast modulations can be resolved by a neuron. One should be aware, though, that the construction principle of MTFs is based on a large amount of averaging, and therefore, the information provided by single spike trains may be lower than that suggested by a MTF (Wohlgemuth et al. 2011).

The second paradigm, gap detection, has been applied in behavioral experiments to grasshoppers and crickets (von Helversen 1972; von Helversen and von Helversen 1997; Schneider and Hennig 2012). Grasshoppers detect gaps of 2–3 ms duration and in this respect are not inferior to vertebrates (Fig. 3b, Prinz and Ronacher 2002). This high resolution of gaps seems to be mediated by the specific interactions between inhibition and

excitation in one identified interneuron (Fig. 3c; Ronacher and Stumpner 1988). The temporal resolution of cricket ears is lower compared to grasshoppers; minimal detectable gap widths are between 6 and 8 ms (Fig. 3d, Schneider and Hennig 2012). One reason for this may stem from the males' sound production system: cricket songs are produced by a resonant mechanism which precludes very fast amplitude changes (Bennet-Clark 1998). Hence, on the receiver's side there is no need to push the temporal resolution to extremes. In addition, compared to broadband signals, which are typical in many grasshoppers and bushcrickets, pure tone signals, such as those produced by many crickets, tend to be more strongly affected by random amplitude fluctuations when traveling through the habitat, which also sets limits for temporal resolution (Römer and Lewald 1992).

Temporal integration refers to the *time-intensity trading* paradigm. In these experiments the minimal audible threshold was found to depend on the duration of the stimuli used. To detect very short stimuli, e.g., of 5 ms duration, higher sound intensities are necessary than for longer, e.g., 100 ms, stimuli. The product of sound intensity and stimulus duration determines the threshold up to durations around 200–300 ms, whereas for longer stimuli, the threshold stays constant (Green 1985). Hence, we are confronted with an apparent discrepancy, the temporal integration-resolution paradox (de Boer 1985): gap detection and MTF paradigms yield time constants in the order of 1–6 ms, whereas from the time-intensity trading paradigm, we are left with time constants in the range of 150–300 ms (for a discussion of this paradox and possible solutions, see de Boer 1985; Viemeister and Wakefield 1991; Tougaard 1998; Pohl et al. 2013).

Transformation of Coding Along the Auditory Pathway

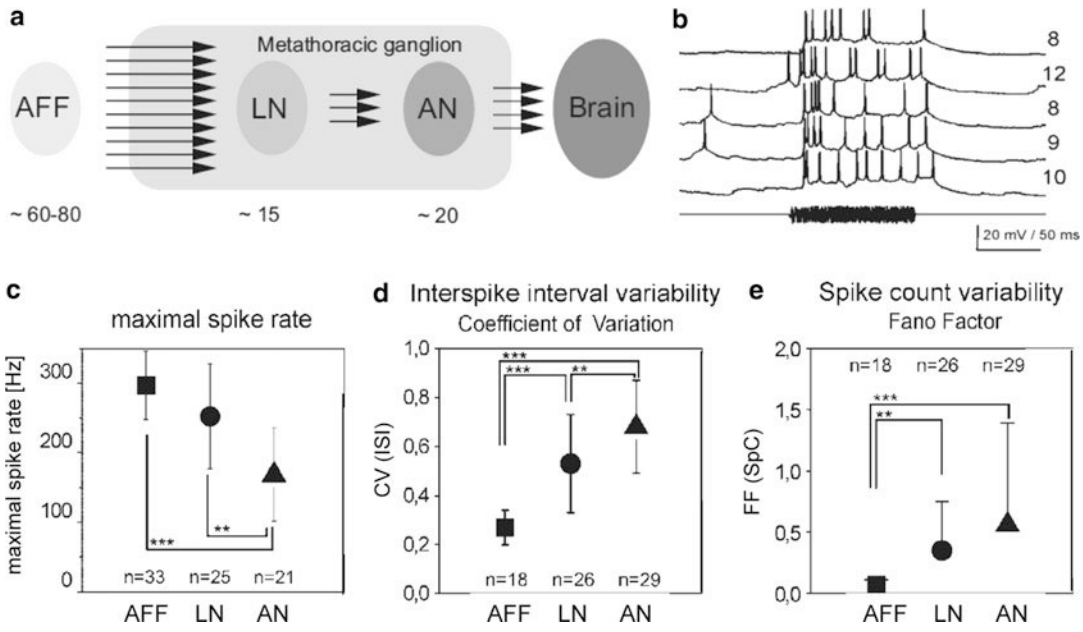
In insects and other arthropods, many neurons can be uniquely identified on the basis of their characteristic morphology. The peripheral stage of a grasshopper's auditory pathway comprises sensory neurons (afferents), local neurons whose processes are confined to the thoracic ganglia, and ascending neurons, whose axons reach the brain. Present knowledge indicates that this corresponds to a feedforward network (Fig. 4a, Vogel et al. 2005; Vogel and Ronacher 2007).

Auditory afferents exhibit high firing rates, up to several hundred Hertz (Fig. 4c). Their tonic spike responses represent the amplitude modulation patterns of auditory stimuli by a modulation of the firing rate. The variability of their responses is rather low (see Fig. 4d, e). The precise responses of sensory neurons allow for a good discrimination and classification of auditory stimulus ensembles (Machens et al. 2003; Wohlgemuth and Ronacher 2007).

Along the auditory pathway the maximal spike rates decrease, whereas the spike train variability increases (both spike count and inter-spike-interval variability, Fig. 4c–e). In accordance with the larger variability of higher-order neurons, the temporal resolution and the classification success for similar stimuli decrease markedly among the ascending neurons. At the level of afferents and among primary-like local neurons, we find a high classification success that depends almost exclusively on the timing of spikes. In contrast, among ascending neurons the classification success based on a single neuron's responses decreases, and spike count differences between stimuli become more important (Wohlgemuth and Ronacher 2007). Among ascending neurons, the information appears to be distributed among several neurons and to be represented as a labeled-line population code (Clemens et al. 2011, 2012). A similar reduction in spike rates from ascending to brain neurons is observed within the auditory pathways of crickets (Schildberger 1984; Kostarakos and Hedwig 2012).

Central Processing in the Frequency or Time Domain?

With Fourier analysis, a signal's temporal structure can be broken down into sine waves, each having a particular amplitude and phase (see, e.g., Yost 2000). If both the resulting spectra for amplitude and phase are known, the original signal can be fully reconstructed. Therefore, there are two principal means of processing a periodic signal: an analysis in the frequency domain, i.e., of the



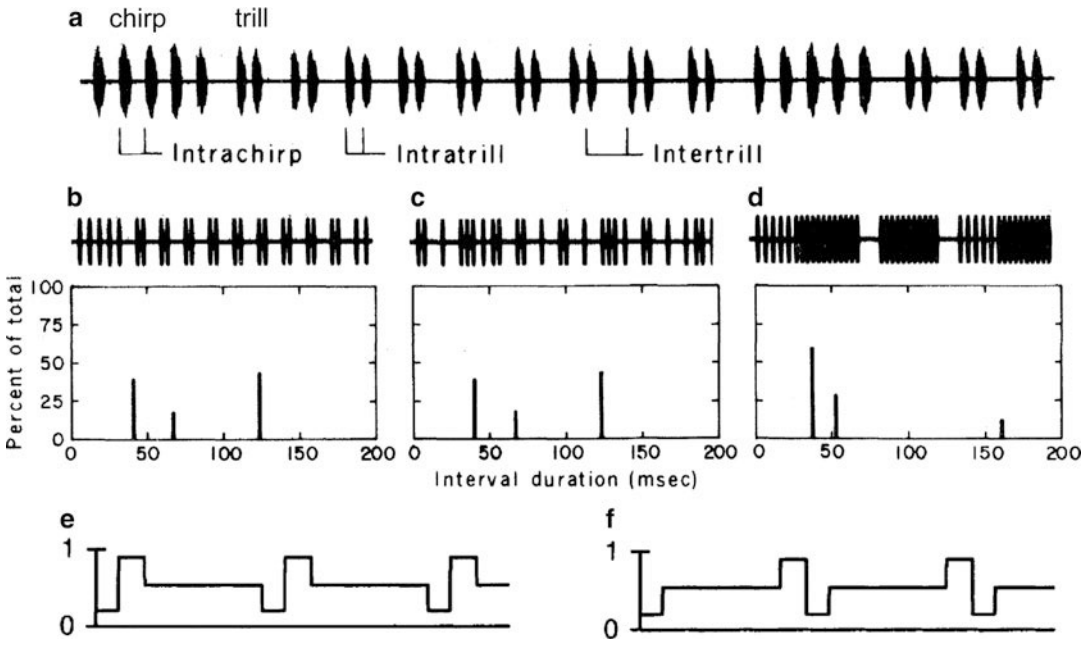
Auditory Processing in Insects, Fig. 4 Transformation of coding along the auditory pathway. (**a**) Scheme of a grasshopper's auditory pathway (AFF auditory afferents, LN local neurons, AN neurons whose axons ascend to the brain). Numbers indicate approximate numbers of neurons at the respective levels. (**b**) Response (numbers of action potentials at right) of an ascending neuron (AN1) to five presentations of an identical stimulus. (**c**) Maximal spike rates at different processing levels. (**d**) Variability of interspike intervals (CV variation coefficient). (**e**) Variability of spike count (expressed as Fano factor FF). Axis in (**c**–**e**) *AFF* afferents, *LN* local neurons, and *AN* ascending neurons; numbers indicate sample size ((**a**) From Ronacher 2013; (**b**–**e**) modified from Vogel et al. 2005, with permission)

amplitude spectrum without phase and thus without temporal information, and an analysis in the time domain by the computation of temporal parameters such as durations and periods of events. For instance, the processing of the sound carrier by a traveling wave is equivalent to the computation of an amplitude spectrum for a frequency analysis.

The envelopes of the communication signals of many insect species have a highly regular and repetitive structure and consist of a series of stereotyped subunits. Remarkably, several experiments have shown that crickets and grasshoppers accept song signals with randomized or shuffled patterns as conspecific (Fig. 5, Pollack and Hoy 1979; von Helversen and von Helversen 1998; Schmidt et al. 2008). This suggested that central processing may be restricted to the frequency domain and serve to compute an amplitude spectrum of the song envelope. A crucial experiment to determine whether a signal is processed in the frequency or time domain is to present reversed or inverted versions of an attractive signal (see Fig. 5e, f). Inverted variants exhibit the same amplitude spectrum as the original but differ in their phase components and thus temporal qualities. If such modified song models show the same attractiveness as the original, a spectral analysis of the signal envelope is very likely as opposed to temporal processing. However, evidence from experiments as shown in Fig. 5e, f and others revealed large differences in attractiveness which suggests that insects process the amplitude modulations of sound stimuli in the time domain (von Helversen and von Helversen 1998; Schmidt et al. 2008; Hennig 2009).

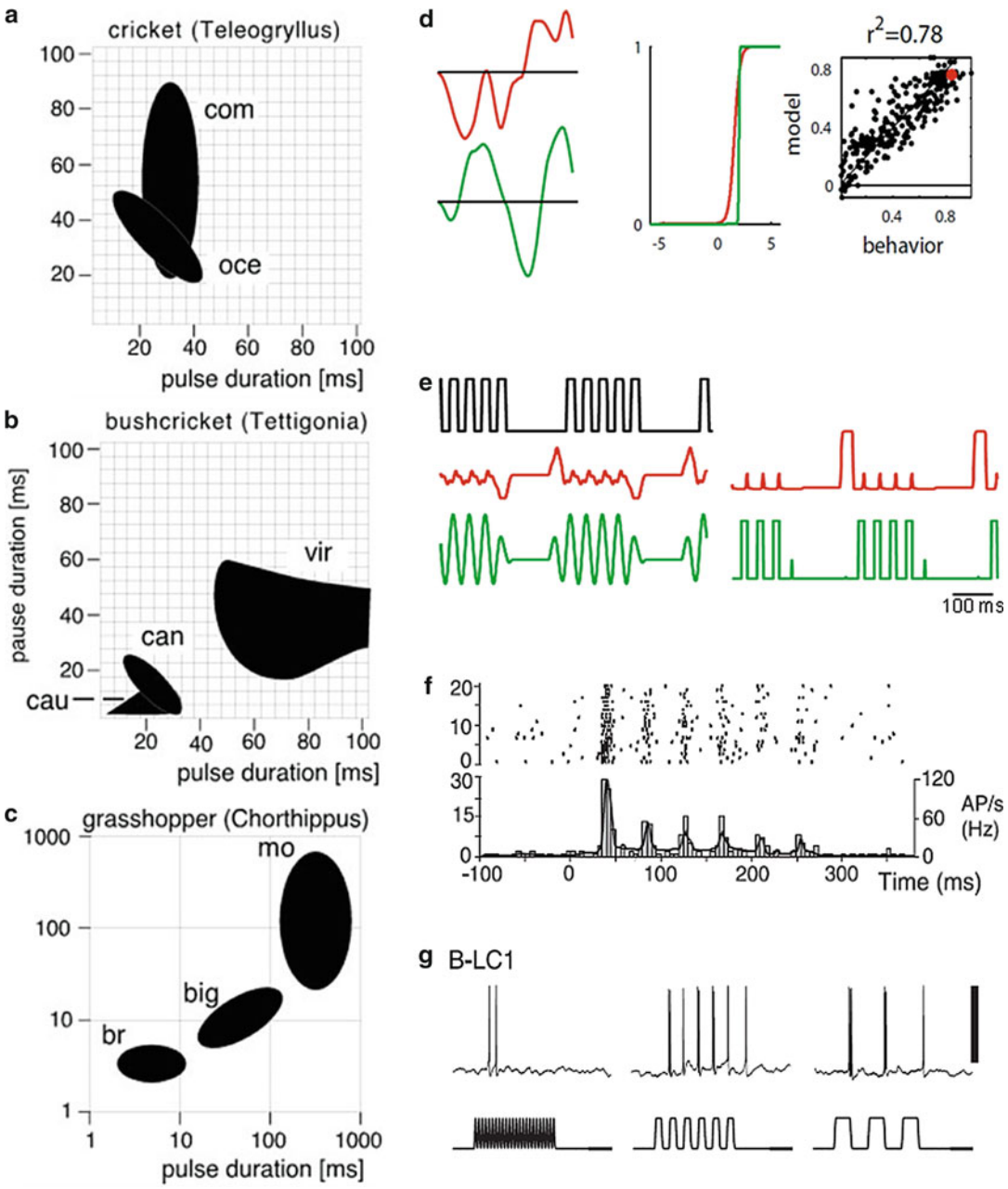
Global Algorithms of Coding

Which global algorithms of coding are implemented in the auditory pathways of insects



Auditory Processing in Insects, Fig. 5 Regular song patterns and the importance of temporal order for recognition. (**a**) Regular calling song of the cricket *Teleogryllus oceanicus*. (**b**–**d**) Song models (*top*) for tests of preference of females and relative frequency of different pulse intervals (*bottom*). Note the differences in the interval distributions between (**b, c**, and **d**). (**b**) Song model of *T. oceanicus*. (**c**) Shuffled song model with the same relative frequency of pulse intervals as in (**a**). (**d**) Song model of *T. commodus*. Song models of (**b**) and (**c**) are attractive for females of *T. oceanicus*, although the temporal order in **c** is randomized. The song pattern in (**d**) from the sibling species is not attractive. Although this experiment suggested central processing of the signal envelope in the frequency domain, the bulk of experimental evidence demonstrates processing in the time domain by crickets and other Insects. (**e**, **f**) Envelopes of reversed song models for behavioral tests with grasshoppers. The pattern in (**e**) is very attractive, but the reversed version in (**f**) is rejected (von Helversen and von Helversen 1998) ((**a**–**d**) From Pollack and Hoy 1979 with permission; (**e**, **f**) from von Helversen and von Helversen 1998, with permission)

to achieve the goals of hearing (see Overview and Background)? In the context of acoustic communication as well as predator avoidance, insects have to recognize specific sound signals for which they possess an innate, internal representation. Presently there is no evidence for a maplike neural representation of specific features in the auditory pathways of insects, except for the tonotopic frequency maps in the periphery (Hildebrandt 2014). For predator avoidance ultrasonic cues and the detection of strong onsets as a typical effect of intense bat calls appear to be most important. Burst coding in the auditory pathways of crickets and grasshoppers shows how onsets are detected (see below, Marsat and Pollack 2006; Creutzig et al. 2009; see also Krahe and Gabbiani 2004). For the recognition of conspecific signals, several global schemes were proposed: autocorrelation, cross-correlation with a template, or combinations of high-, low-, and band-pass filters. All these schemes are derived from Fourier transformation and represent mathematical and technical solutions, for which evidence from several investigations exists (Schildberger 1994; Weber and Thorson 1989; Hennig 2003). However, only a few studies have been able to reproduce the preferences exhibited by female insects over a wider range of signals (see Fig. 6a–c for response profiles from several species of crickets, bushcrickets, and grasshoppers upon presentation of stereotyped sound patterns composed of subunits built by regular pulses and pauses (von Helversen and von Helversen 1994)). In a recent alternative approach, the recognition of sound signals was examined using linear-nonlinear (LN) models adapted from computational



Auditory Processing in Insects, Fig. 6 Preference profiles of female insects for song signals and LN models. (**a**–**c**) Preference profiles for song signals in the time domain by crickets, bush crickets, and grasshoppers (A: oce oceanicus, com commodus; B: cau caudata, can cantans, vir viridissima; C: br brunneus, big biguttulus, mo mollis). (**d**–**e**): LN models account for preference functions of the cricket *Gryllus bimaculatus*. (**d**) Two LN models (*red*, *green*) that predict the preferences for pulse patterns by female crickets (*G. bimaculatus*): linear filters (*left panels*) and the respective nonlinearities (*central panel*) that predict behavioral responses of female crickets (*right panel*). Note that the duration of the linear filters is only 64 ms. The *red dot* in the *right panel* of (**d**) refers to the pattern in (**e**). (**e**) Output of the LN models (**d**) in response to a song model (*upper trace*). Response of the linear filters to the song model before (*left panels*) and after passing through the nonlinearity (*right panels*). The upper filter (*red* in **d** and **e**) responds like an onset detector; the lower filter resembles a Gabor-function and selectively responds to

neuroscience (Clemens and Hennig 2013; Clemens and Ronacher 2013). For crickets as well as grasshoppers, Gabor functions emerged as filters that responded best to particular subunit shapes common to the conspecific song signal (combinations of pulse and pause or pairs of pulses, Fig. 6d, e). Gabor functions (a sine wave multiplied with a Gaussian function) are well known from sensory pathways in vertebrates (visual, Daugman 1984; Simoncelli and Olshausen 2001; Priebe and Ferster 2012; auditory, Smith and Lewicki 2006). However, the most notable property of the proposed LN model was the independence of the computation from the exact timing of occurrence of the template within a larger time window. Therefore, this LN-model approach offers an elegant solution to the attractiveness of shuffled and irregular song signals (see Fig. 5 and subchapter: processing in the frequency or time domain). Small modifications of Gabor functions are capable of reproducing response profiles known from several species of insects (Fig. 6a–c, Clemens and Hennig 2013). Although these filters by default describe the output of the whole recognition system, present evidence suggests that certain identified neurons in crickets may represent a neuronal correlate of specific LN features (compare Fig. 6e–g, Zorović and Hedwig 2011; Kostarakos and Hedwig 2012; Clemens and Hennig 2013).

Local Mechanisms of Coding

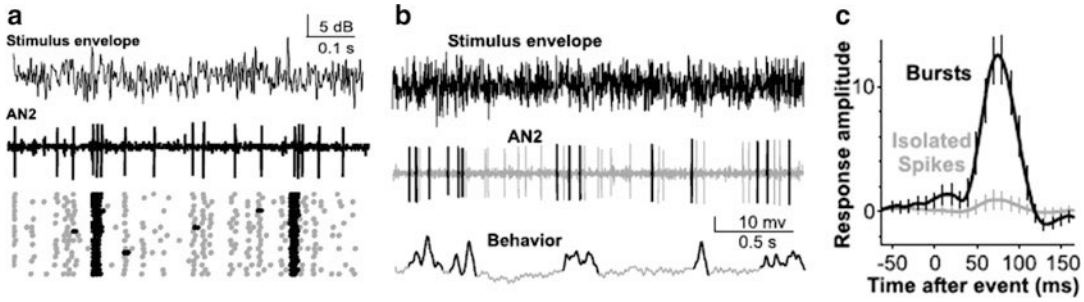
For insects there exist several prominent examples of how global algorithms of coding are at least in part implemented by identified neurons. For instance, the detection of bat predators by crickets is mediated by a specific auditory neuron (AN2), whose bursting is crucial for a behavioral response (Fig. 7, Nolen and Hoy 1984; Marsat and Pollack 2006). Similarly, the AN12 in grasshoppers codes for a specific song feature by bursts (Creutzig et al. 2009). Specific combinations of excitation and inhibition account for feature extraction in grasshoppers (gap detection by AN4; Ronacher and Stumpner 1988 – see temporal resolution) and serve as a basis for pulse rate detection by the identified neuron B-LI4 in crickets (Kostarakos and Hedwig 2012). Similarly, resonant properties within the auditory pathway of bushcrickets mediate the detection of specific pulse rates and can be modeled as a property of single neurons (Bush and Schul 2006; Webb et al. 2007). Short time constants of single identified neurons allow them to act as feature detectors for ultrasonic pulses in the auditory pathway of moths (Boyan and Fullard 1988).

Processing under the Constraints of Noise: A Twofold Problem

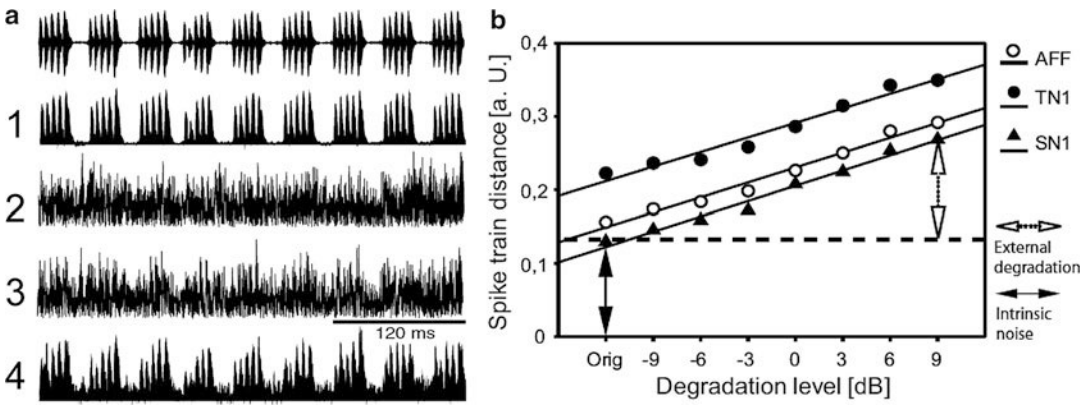
Noise poses unavoidable problems for all sensory systems. There are two classes of noise, external and internal. Various types of external noise influence sound waves on their way from sender to receiver. Hence, as a rule, receivers have to cope with signals that are masked and degraded in their temporal structure (e.g., Michelsen and Larsen 1983; Römer et al. 1989; Römer 2001; Schmidt et al. 2011; see also Brumm and Slabbekoorn 2005; Wiley 2006).

Crickets with their pure tone songs have found a solution to reduce the impact of external noise. The females' hearing system is tuned to the carrier frequency of male songs and becomes increasingly less sensitive to frequencies farther from the carrier frequency (Fig. 2g, h). The sharpness of the tuning curve may depend on ecological conditions (Schmidt et al. 2011). A cricket species (*Paroecanthus podagrosus*) living in very noisy tropical rainforests exhibits an exceptionally narrow tuning ($Q_{10dB} \sim 4$) that allows for an efficient

Auditory Processing in Insects, Fig. 6 (continued) pairs of pulses (*green* in **d** and **e**). (**f–g**) The response pattern of an auditory interneuron in the cricket brain resembles the output of the onset detector (*red* in **d**, **e**). The neuronal discharges in (**g**) illustrate onset responses to sound patterns with different pulse rates ((**a–c**) From Hennig et al. 2004; (**d**, **e**) modified from Clemens and Hennig 2013; and (**f**, **g**) from Zorović and Hedwig 2011, with permission)



Auditory Processing in Insects, Fig. 7 Burst coding by an identified interneuron (AN2) and avoidance behavior in crickets to bat like stimuli. (**a**) Bursts in AN2 in response to amplitude-modulated stimuli (carrier frequency: 30 kHz); burst spikes are marked as *black* dots in raster plot. (**b**) Response of AN2 to a sound stimulus (as in **a**) and behavioral response (*bottom trace*: abdominal movements away from the sound source). (**c**) Amplitude of abdominal movements after isolated action potentials (*gray*) or bursts (*black*). Positive values indicate abdomen flexion away from the sound source (From Marsat and Pollack 2006, with permission)



Auditory Processing in Insects, Fig. 8 Processing under the constraints of noise: a twofold problem. (**a**) Song of a tropical cricket *Paroecanthus podagrosus* (*upper trace*). Lower traces. 1: Song envelope. 2: Song envelope under ambient noise levels as recorded in the habitat. 3: Song plus noise less efficiently filtered with the broader tuning curve of a European cricket (*Gryllus bimaculatus*). 4: Song plus noise filtered with the narrow tuning curve of *P. podagrosus*, compare Fig. 2h (Adapted from Schmidt et al. 2011). (**b**) Effects of intrinsic noise and external signal degradation on spike train dissimilarities of three representative neurons, assessed with the van Rossum metric and corrected for spike rate differences. The *black arrow* at "orig" indicates the average spike train distance found for repeated presentation of the original song pattern, i.e., the result of trial-to-trial variability. The *open arrow* indicates the additional distance caused by the most strongly degraded signal. AFF sensory neuron, TN1, and SN1 two primary-like local neurons. ((**a**) Modified from Schmidt et al. 2011. (**b**) Modified from Neuhofer et al. 2011 and Ronacher 2013, with permission)

suppression of ambient noise (Figs. 2h, 8a). Thus, in this species we find a narrow peripheral filter that is perfectly matched to the carrier frequency of conspecific signals and dismisses signals from other species, at the expense of reducing the range of perceivable sounds (Fig. 8a). Many bushcrickets use broadband communication signals which yield a twofold advantage. First, broadband signals are less likely to be degraded in the biotope compared to narrowband signals (Michelsen and Larsen 1983; Römer and Lewald 1992). In addition, the receiver's nervous system can compare the neuronal signals from differently tuned auditory receptors and thereby reduce intrinsic neuronal noise (see below). Note that the hearing systems of vertebrates (mammals,

owls) also use broadband signals to reduce noise or to resolve localization ambiguities (Konishi 1990).

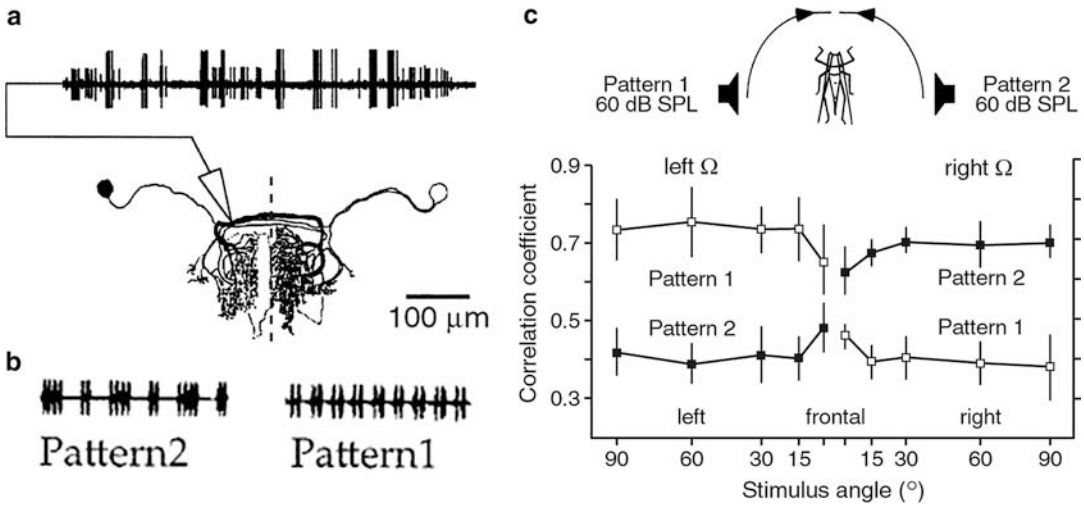
In addition to external noise, auditory systems face a second noise problem. Neuronal signals are inherently noisy due to the stochastic opening and closing of ion channels. This *intrinsic noise* becomes evident as trial-to-trial variability of spike trains in response to repeated presentations of an identical stimulus (see Fig. 4b). Animals with a large nervous system may alleviate this problem by averaging responses from many neurons with similar properties. However, due to size constraints insects probably cannot afford this solution. Under some conditions noise may play a beneficial role and improve neural computations, for example, by stochastic resonance (for review see McDonnell and Ward 2011).

To quantify the trial-to-trial variability or to compare neuronal responses to different stimuli, one can apply a *spike train metric* (see, e.g., van Rossum 2001). This metric describes the similarity of two spike trains by a single number, which is in line with our intuition of distance: small values indicate high similarity. The metric uses an adjustable parameter to take into account both differences in the spike count as well as in the timing of spikes. This metric approach was used to determine the relative impacts of external and intrinsic noise on the encoding of envelope-degraded stimuli by auditory neurons of grasshoppers (Neuhofer et al. 2011). Unexpectedly, the contribution of external signal degradation to the overall spike train distances was low: even for the highest degradation level, its amount did not exceed that of intrinsic noise (Fig. 8b).

As long as rather different signals have to be analyzed, neuronal noise may not be a very serious problem. However, with communication signals that serve to attract mates, it will be necessary to discriminate between similar signals and to detect small deviations from a species-specific pattern. This is especially the case if quality cues from the sound signals of potential partners are to be extracted. Remarkably enough, the spike trains of even a single auditory afferent allow for an almost perfect discrimination of songs of different males of one species (Machens et al. 2003). At higher stages in the auditory pathway, however, the discrimination deteriorates and intrinsic noise cannot be neglected as a limiting factor. Grasshoppers appear to have circumvented this intrinsic noise problem by changing the coding scheme at a rather peripheral stage of processing (see Fig. 4a): among the ascending neurons information is distributed according to a labeled-line code (Clemens et al. 2011, 2012; see also Clemens and Ronacher 2013).

Directional Hearing

The biophysical qualities of the pressure difference receivers equip the ears of insects with a directional dependence of the vibrational amplitudes of their tympana (Michelsen 1998). Sensory neurons reflect this dependence in spike numbers, and for some species, this difference in response strength is also translated into timing differences (enhanced by ramps Krahe and Ronacher 1993; Ronacher and Krahe 2000). The contrast in response magnitude of sensory neurons for left and right differences is enlarged by contralateral inhibition from local interneurons (ON neurons in crickets, Selverston et al. 1985; bushcrickets, Römer and Krusch 2000; LN in grasshoppers, Marquart 1985). The representation of sound from one side is therefore enhanced, because the responses to sound from the opposite side are suppressed, and this takes place already at early levels of auditory processing of sound direction, usually at the first synapse after sensory receptors. This mechanism of contralateral inhibition in auditory pathways transforms the acoustic environment of insects into acoustic hemispheres. These hemispheres enable a rather accurate distinction of left and right sound signals at the cost of accuracy in determining the angle of the sound source (lateralization in grasshoppers, von Helversen and Rheinlaender 1988). A corollary of the computation of acoustic hemispheres is that sound signals are selectively represented on one side of the insect or the other. This computation resembles the phenomenon of selective attention



Auditory Processing in Insects, Fig. 9 Acoustic hemispheres and selective representation of sound patterns in bush crickets. (**a**, **b**) *Left* and *right* specimens of the local interneuron ON1 were recorded simultaneously, while different sound signals (pattern 1 and 2) were presented from either side. (**c**) Both interneurons selectively represent only the sound pattern from one side (inset: experimental arrangement, correlation coefficient of the spike train with the sound pattern as a measure for copying fidelity) (From Römer and Krusch 2000, with permission)

(Pollack 1988) and allows insects to discriminate sound sources in much the same way as the cocktail party effect well known from humans (Fig. 9, Römer and Krusch 2000). Notably, contralateral inhibition and acoustic hemispheres can exert a considerable influence on mate choice as females of bushcrickets prefer males that take a leader role among singing males, since the representation of the song signal of a follower is suppressed in their auditory pathway (Hartbauer et al. 2005; Siegert et al. 2011). These local computations for enhanced directional responses can also affect the singing behavior of males and promote synchronization and thus chorusing among males (Greenfield and Roizen 1993; Greenfield 1994; Hartbauer et al. 2005). Due to the computational conflict between representation of sound pattern and sound source, grasshoppers split the sensory pathways for processing of cues for pattern and direction by parallel processing (von Helversen 1984; Ronacher et al. 1986). However, in the evolutionarily older communication systems of crickets and bushcrickets, serial processing of pattern and direction appears to prevail (Wendler 1989; Stabel et al. 1989; von Helversen and von Helversen 1995; Schul et al. 1998).

Conclusions

Auditory processing in insects is constrained by small body size and a relatively small number of neurons. Thus, the hearing capacities are focused on highly relevant tasks, such as predator detection, mate attraction, and, in some cases, eavesdropping on prey or host signals. Insect ears evolved from cuticular mechanosensory proprioreceptors and may respond to different aspects of sound waves – particle velocity or sound pressure; ears can be found in almost any part of the body, and as a consequence, different neural structures are involved in auditory processing. In the auditory pathways, we find similar computational principles and transformations of sensory representations as in vertebrates, however, with an important difference: complex computations are often performed by single neurons or very small populations of identifiable neurons. This size constraint and the focus on a few relevant tasks facilitate experimental approaches.

Acknowledgments We want to thank the members of our lab who contributed to several of the cited studies. Special thanks are due to Dr. Michael Reichert who gave helpful advice on the English style and substantially improved the manuscript, as well as to an anonymous reviewer whom we owe many helpful suggestions.

References

Andersson M, Simmons LW (2006) Sexual selection and mate choice. Trends Ecol Evol 21:296–302

Autrum HJ (1942) Schallempfang bei Mensch und Tier. Naturwissenschaften 5:69–85

Barth FG (2002) A spider's world: senses and behavior. Springer, Berlin/Heidelberg/New York

Benda J, Hennig RM (2008) Spike-frequency adaptation generates intensity invariance in a primary auditory interneuron. J Comput Neurosci 24:113–136

Benda J, Herz AVM (2003) A universal model for spike-frequency adaptation. Neural Comput 15:2523–2564

Bennet-Clark HC (1998) Size and scale effects as constraints in insect sound communication. Phil Trans R Soc Lond B 353:407–419

Bentley DR, Hoy RR (1972) The genetic control of the neuronal network generating cricket (*Teleogryllus gryllus*) song pattern. Anim Behav 20:478–492

Bernal XE, Rand AS, Ryan MJ (2006) Acoustic preferences and localization performance of blood-sucking flies (*Corethrella Coquillett*) to tungara frog calls. Behav Ecol 17:709–715

Boyan GS, Fullard JH (1988) Information processing at a central synapse suggests a noise filter in the auditory pathway of the noctuid moth. J Comp Physiol A 164: 251–258

Brown CH (1994) Sound localization. In: Fay RR, Popper AN (eds) Comparative hearing: mammals. Springer, New York/Berlin, pp 57–96

Brumm H, Slabbekoorn H (2005) Acoustic communication in noise. Adv Stud Behav 35:151–209

Bura VL, Rower VG, Martin PR, Yack JE (2011) Whistling in caterpillars (*Amorpha juglandis*, Bombycoidea): sound-producing mechanism and function. J Exp Biol 214:30–37

Bush SL, Schul J (2006) Pulse-rate recognition in an insect: evidence of a role for oscillatory neurons. J Comp Physiol A 192:113–121

Clemens J, Hennig RM (2013) Computational principles underlying the recognition of acoustic signals in insects. J Comput Neurosci 35:75–85

Clemens J, Ronacher B (2013) Feature extraction and integration underlying perceptual decision making during courtship in grasshoppers. J Neurosci 33: 12136–12145

Clemens J, Kutzki O, Ronacher B, Schreiber S, Wohlgemuth S (2011) Efficient transformation of an auditory population code in a small sensory system. Proc Natl Acad Sci U S A 108:13812–13817

Clemens J, Wohlgemuth S, Ronacher B (2012) Nonlinear computations underlying temporal and population sparseness in the auditory system of the grasshopper. J Neurosci 32:10053–10062

Creutzig F, Wohlgemuth S, Stumpner A, Benda J, Ronacher B, Herz AVM (2009) Time-scale invariant representation of acoustic communication signals by a bursting neuron. J Neurosci 29:2575–2580

Daugman JG (1984) Spatial visual channels in the Fourier plane. Vis Res 24:891–910

de Boer E (1985) Auditory time constants: a paradox? In: Michelsen A (ed) Time resolution in auditory systems. Springer, Berlin/Heidelberg, pp 141–157

Faure PA, Mason AC, Yack JE (2009) Invertebrate ears and hearing. In: Binder MD, Hirokawa N, Windhorst U, Hirsch MC (eds) Encyclopedia of neuroscience. Springer, Berlin, pp 2035–2042

Franz A, Ronacher B (2002) Temperature dependence of temporal resolution in an insect nervous system. J Comp Physiol A 188:261–271

Fullard JH, Yack JE (1993) The evolutionary biology of insect hearing. Trends Ecol Evol 8:248–252

Gerhardt HC, Huber F (2002) Acoustic communication in insects and anurans. University of Chicago Press, Chicago

Gollisch T, Herz AVM (2004) Input-driven components of spike-frequency adaptation can be unmasked in vivo. J Neurosci 24:7435–7444

Gollisch T, Herz AVM (2005) Disentangling sub-millisecond processes within an auditory transduction chain. PLoS Biol 3(e8):1–11

Green DM (1985) Temporal factors in psychoacoustics. In: Michelsen A (ed) Time resolution in auditory systems. Springer, Berlin/Heidelberg, pp 122–140

Greenfield MD (1994) Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions. Am Zool 34:605–615

Greenfield MD, Roizen I (1993) Katydid synchronous chorusing is an evolutionary stable outcome of female choice. Nature 364:618–620

Grothe B (2000) The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. Prog Neurobiol 61:581–610

Hartbauer M, Kratzer S, Steiner K, Römer H (2005) Mechanisms for synchrony and alternation in song interactions of the bushcricket *Mecopoda elongata* (Tettigoniidae, Orthoptera). J Comp Physiol A 191:175–188

Hennig RM (2003) Acoustic feature extraction by cross-correlation in crickets? J Comp Physiol A 189:589–598

Hennig RM (2009) Walking in Fourier's space: algorithms for the computation of periodicities in song patterns by the cricket *Gryllus bimaculatus*. J Comp Physiol A 195:971–987

Hennig RM, Franz A, Stumpner A (2004) Processing of auditory information in insects. Microsc Res Tech 63: 351–374

Hildebrandt KJ (2014) Neural maps in insect versus vertebrate auditory systems. Curr Opin Neurobiol 24:82–87

Hildebrandt KJ, Benda J, Hennig RM (2009) The origin of adaptation in the auditory pathway of locusts is specific to cell type and function. J Neurosci 29:2626–2636

Hildebrandt KJ, Benda J, Hennig RM (2011) Multiple arithmetic operations in a single neuron: the recruitment of adaptation processes in the cricket auditory pathway depends on sensory context. J Neurosci 31: 14142–14150

Hoy RR (1989) Startle, categorical response, and attention in acoustic behavior of insects. Ann Rev Neurosci 12: 355–375

Hoy RR, Popper AN, Fay RR (eds) (1998) Comparative hearing: insects. Springer, New York

Huber F (1992) Verhalten und Neurobiologie von stimmbegabten Insekten. Naturwissenschaften 79: 393–406

Huber F, Markl H (1983) Neuroethology and behavioural physiology: roots and growing points. Springer, Heidelberg/New York

Huber F, Moore TE, Loher W (eds) (1989) Cricket behavior and neurobiology. Cornell University Press, Ithaca

Hummel J, Kössl M, Nowotny M (2011) Sound-induced tympanal membrane motion in bushcrickets and its relationship to sensory output. J Exp Biol 214: 3596–3604

Janssen R (1992) Thermal influences on nervous system function. Neurosci Biobehav Rev 16:399–413

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84: 541–577

Konishi M (1990) Similar algorithms in different sensory systems and animals. Cold Spring Harb Symp Quant Biol 55:575–584

Kostarakos K, Hedwig B (2012) Calling song recognition in female crickets: temporal tuning of identified brain neurons matches behavior. J Neurosci 32:9601–9612

Krahe R, Gabbiani F (2004) Burst firing in sensory systems. Nat Rev Neurosci 5:13–24

Krahe R, Ronacher B (1993) Long rise times of sound pulses in grasshopper songs improve the directionality cues received by the CNS from auditory receptors. J Comp Physiol A 173:425–434

Lakes-Harlan R, Stölting H, Stumpner A (1999) Convergent evolution of insect hearing organs from a preadaptive structure. Proc R Soc Lond B 266: 1161–1167

Lehmann GUC, Strauß J, Lakes-Harlan R (2007) Listening when there is no sexual signalling? Maintenance of hearing in the asexual bushcricket *Poecilimon intermedius*. J Comp Physiol A 193:537–545

Machens CK, Stemmler MB, Prinz P, Krahe R, Ronacher B, Herz AVM (2001) Representation of acoustic communication signals by insect auditory receptor neurons. J Neurosci 21:3215–3227

Machens CK, Schütze H, Franz A, Stemmler MB, Ronacher B, Herz AVM (2003) Auditory receptor neurons preserve characteristic differences between conspecific communication signals. Nat Neurosci 6: 341–342

Marquart V (1985) Local interneurons mediating excitation and inhibition onto ascending neurons in the auditory pathway of grasshoppers. Naturwissenschaften 72: 42–44

Marsat G, Pollack GS (2006) A behavioural role for feature detection by sensory bursts. J Neurosci 26: 10542–10547

McDonnell MD, Ward LM (2011) The benefits of noise in neural systems: bridging theory and experiment. Nat Rev Neurosci 12:415–425

Meier T, Reichert H (1990) Embryonic development and evolutionary origin of the orthopteran auditory organs. J Neurobiol 21:592–610

Michelsen A (1979) Insect ears as mechanical systems. Am Sci 67:696–706

Michelsen A (ed) (1985) Time resolution in auditory systems. Springer, Berlin/Heidelberg

Michelsen A (1998) Biophysics of sound localization in insects. In: Hoy RR, Popper AN, Fay RR (eds) Comparative hearing: insects. Springer, Berlin/New York, pp 18–62

Michelsen A, Larsen ON (1983) Strategies for acoustic communication in complex environments. In: Huber F, Markl H (eds) Neuroethology and behavioural physiology. Springer, Berlin/Heidelberg, pp 321–331

Michelsen A, Popov A, Lewis B (1994) Physics of directional hearing in the cricket *Gryllus bimaculatus*. J Comp Physiol A 175:153–164

Miller LA, Surlykke A (2001) How some insects detect and avoid being eaten by bats: tactics and countertactics of prey and predator. Bioscience 51:571–582

Moiseff A, Pollack G, Hoy R (1978) Steering responses of flying crickets to sound and ultrasound: mate attraction and predator avoidance. Proc Natl Acad Sci U S A 75: 4052–4056

Montealegre-Z F, Jonsson T, Robert D (2011) Sound radiation and wing mechanics in stridulating field crickets (Orthoptera: Gryllidae). J Exp Biol 214: 2105–2117

Montealegre-Z F, Jonsson T, Robson-Brown KA, Postles M, Robert D (2012) Convergent evolution between insect and mammalian audition. Science 338: 968–971

Nadrowski B, Effertz T, Senthilan PR, Göpfert MC (2011) Antennal hearing in insects – new findings, new questions. Hearing Res 273:7–13

Neuhofer D, Stemmler M, Ronacher B (2011) Neuronal precision and the limits for acoustic signal recognition in a small neuronal network. J Comp Physiol A 197: 251–265

Nolen TG, Hoy RR (1984) Initiation of behavior by single neurons: the role of behavioral context. Science 226: 992–994

Penzlin H (2005) Lehrbuch der Tierphysiologie. Elsevier, München

Pohl NU, Slabbekoorn H, Neubauer H, Heil P, Klump GM, Langemann U (2013) Why longer song elements are easier to detect: threshold level-duration functions in the great tit and comparison with human data. J Comp Physiol A 199:239–252

Pollack GS (1988) Selective attention in an insect auditory neuron. J Neurosci 8:2635–2639

Pollack GS, Hoy RR (1979) Temporal pattern as a cue for species-specific calling song recognition in crickets. Science 204:429–432

Priebe NJ, Ferster D (2012) Mechanisms of neuronal computation in mammalian visual cortex. Neuron 75: 194–208

Prinz P, Ronacher B (2002) Temporal modulation transfer functions in auditory receptor fibres of the locust (*Locusta migratoria* L.). J Comp Physiol A 188: 577–587

Riede K (1987) A comparative study of mating behaviour in some neotropical grasshoppers (Acridoidea). Ethology 76:265–296

Riede K, Kämper G, Höfler I (1990) Tympana, auditory thresholds, and projection areas of tympanal nerves in singing and silent grasshoppers (insects, Acridoidea). Zoomorphology 109:223–230

Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W (1997) Spikes – exploring the neural code. MIT Press, Cambridge, MA

Robert D, Göpfert MC (2002) Novel schemes for hearing and orientation in insects. Curr Opin Neurobiol 12: 715–720

Robert D, Hoy RR (1998) The evolutionary innovation of tympanal hearing in Diptera. In: Hoy RR, Popper AN, Fay RR (eds) Comparative hearing: insects. Springer, New York, pp 197–227

Robert D, Miles RN, Hoy RR (1996) Directional hearing by mechanical coupling in the parasitoid fly *Ormia ochracea*. J Comp Physiol A 179:29–44

Robertson RM, Money TG (2012) Temperature and neuronal circuit function: compensation, tuning and tolerance. Curr Opin Neurobiol 22:724–734

Römer H (1976) Die Informationsverarbeitung tympanaler Rezeptorelemente von *Locusta migratoria*. J Comp Physiol A 109:101–122

Römer H (1983) Tonotopic organization of the auditory neuropile in the bushcricket *Tettigonia viridissima*. Nature 306:60–62

Römer H (2001) Ecological constraints for sound communication: from grasshoppers to elephants. In: Barth FG, Schmid A (eds) Ecology of sensing. Springer, Berlin/Heidelberg/New York, pp 59–77

Römer H, Krusch M (2000) A gain-control mechanism for processing of chorus sounds in the afferent auditory pathway of the bushcricket *Tettigonia viridissima* (Orthoptera, Tettigoniidae). J Comp Physiol A 186: 181–191

Römer H, Lewald J (1992) High-frequency sound transmission in natural habitats: implications for the evolution of insect acoustic communication. Behav Ecol Sociobiol 29:437–444

Römer H, Marquart V, Hardt M (1988) Organization of a sensory neuropile in the auditory pathway of two groups of Orthoptera. J Comp Neurol 275:201–215

Römer H, Bailey WJ, Dadour I (1989) Insect hearing in the field: III masking by noise. J Comp Physiol A 164: 609–620

Römer H, Spickermann M, Bailey W (1998) Sensory basis for sound intensity discrimination in the bushcricket *Requena verticalis* (Tettigoniidae, Orthoptera). J Comp Physiol A 182:595–607

Ronacher B (2013) Processing of species-specific signals in the auditory pathway of grasshoppers. In: Hedwig B (ed) Insect hearing and acoustic communication. Springer, Berlin, Heidelberg, pp 185–204

Ronacher B, Krahe R (2000) Temporal integration vs. parallel processing: coping with the variability of neuronal messages in directional hearing of insects. Eur J Neurosci 12:2147–2156

Ronacher B, Stumpner A (1988) Filtering of behaviourally relevant temporal parameters of a grasshopper's song by an auditory interneuron. J Comp Physiol A 163: 517–523

Ronacher B, von Helversen D, von Helversen O (1986) Routes and stations in the processing of auditory directional information in the CNS of a grasshopper, as revealed by surgical experiments. J Comp Physiol A 158:363–374

Ronacher B, Franz A, Wohlgemuth S, Hennig H (2004) Variability of spike trains and the processing of temporal patterns of acoustic signals–problems, constraints, and solutions. J Comp Physiol A 190: 257–277

Schildberger K (1984) Temporal selectivity of identified auditory neurons in the cricket brain. J Comp Physiol A 155:171–185

Schildberger K (1994) The auditory pathway of crickets: adaptations for intraspecific acoustic communication. In: Schildberger K, Elsner N (eds) Neural basis of behavioural adaptations. G Fischer, Stuttgart, pp 209–225

Schildberger K, Elsner N (1994) Neural basis of behavioural adaptations. G. Fischer, Stuttgart

Schmidt A, Ronacher B, Hennig RM (2008) The role of frequency, phase and time for processing amplitude modulated signals by grasshoppers. J Comp Physiol A 194:221–233

Schmidt AKD, Riede K, Römer H (2011) High background noise shapes selective auditory filters in a tropical cricket. J Exp Biol 214:1754–1762

Schneider E, Hennig RM (2012) Temporal resolution for calling song signals by female crickets, *Gryllus bimaculatus*. J Comp Physiol A 198:181–191

Schul J, Sheridan RA (2006) Auditory stream segregation in an insect. Neuroscience 138:1–4

Schul J, von Helversen D, Weber T (1998) Selective phonotaxis in *Tettigonia cantans* and *T. viridissima* in song recognition and discrimination. J Comp Physiol A182:687–694

Selverston A, Kleindienst H-U, Huber F (1985) Synaptic connectivity between cricket auditory interneurons as studied by selective photoinactivation. J Neurosci 5: 1283–1292

Senthilan PR, Piepenbrock D, Ovezmyradov G, Nadrowski B, Bechstedt S, Pauls S, Winkler M, Möbius W, Howard J, Göpfert MC (2012) Drosophila auditory organ genes and genetic hearing defects. Cell 150:1042–1054

Siegert ME, Römer H, Hashim R, Hartbauer M (2011) Neuronal correlates of a preference for leading signals

in the synchronizing bushcricket *Mecopoda elongata* (Orthoptera: Tettigoniidae). J Exp Biol 214:3924–3934
Simoncelli E, Olshausen B (2001) Natural image statistics and neural representation. Annu Rev Neurosci 24: 1193–1216
Smith EC, Lewicki MS (2006) Efficient auditory coding. Nature 439:978–982
Stabel J, Wendler G, Scharstein H (1989) Cricket phonotaxis: localization depends on recognition of the calling song pattern. J Comp Physiol A 165:165–177
Stölting H, Stumpner A (1998) Tonotopic organization of auditory receptors of the bushcricket *Pholidoptera griseoaptera* (Tettigoniidae, Decticinae). Cell Tissue Res 294:377–386
Stumpner A (1996) Tonotopic organization of the hearing organ in a bushcricket. Naturwissenschaften 83:81–84
Stumpner A (1997) An auditory interneurone tuned to the male song frequency in the duetting bushcricket *Ancistrura nigrovittata* (Orthoptera, Phaneropteridae). J Exp Biol 200:1089–1101
Stumpner A, Lakes-Harlan R (1996) Auditory interneurons in a hearing fly (*Therobia leonidei*, Ormiini, Tachinidae, Diptera). J Comp Physiol A 178:227–233
Stumpner A, von Helversen D (2001) Evolution and function of auditory systems in insects. Naturwissenschaften 88:159–170
Suga N, Zhang Y, Yan J (1997) Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. J Neurophysiol 77:2098–2114
Tougaard J (1998) Detection of short pure-tone stimuli in the noctuid ear: what are temporal integration and integration time all about? J Comp Physiol A 183:563–572
van Rossum MCW (2001) A novel spike distance. Neural Comput 13:751–763
van Staaden MJ, Römer H (1998) Evolutionary transition from stretch to hearing organs in ancient grasshoppers. Nature 394:773–776
Viemeister NF, Plack CJ (1993) Time analysis. In: Yost WA, Popper AN, Fay RR (eds) Human psychophysics. Springer, Berlin/Heidelberg/New York, pp 116–154
Viemeister NF, Wakefield GH (1991) Temporal integration and multiple looks. J Acoust Soc Am 90:858–865
Vogel A, Ronacher B (2007) Neural correlations increase between consecutive processing levels in the auditory system of locusts. J Neurophysiol 97:3376–3385
Vogel A, Hennig RM, Ronacher B (2005) Increase of neuronal response variability at higher processing levels as revealed by simultaneous recordings. J Neurophysiol 93:3548–3559
von Helversen D (1972) Gesang des Männchens und Lautschema des Weibchens bei der Feldheuschrecke *Chorthippus biguttulus* (Orthoptera, Acrididae). J Comp Physiol 81:381–422
von Helversen D (1984) Parallel processing in auditory pattern recognition and directional analysis by the grasshopper *Chorthippus biguttulus* L (Acrididae). J Comp Physiol A 154:837–846
von Helversen D (1997) Acoustic communication and orientation in grasshoppers. In: Lehrer M (ed) Orientation and communication in arthropods. Birkhäuser, Basel, pp 301–341
von Helversen D, Rheinlaender (1988) Interaural intensity and time discrimination in an unrestrained grasshopper: a tentative behavioural approach. J Comp Physiol A 162:333–340
von Helversen D, von Helversen O (1975a) Verhaltensgenetische Untersuchungen am akustischen Kommunikationssystem der Feldheuschrecken (Orthoptera, Acrididae). I. Der Gesang von Artbastarden zwischen *Chorthippus biguttulus* und *C. mollis*. J Comp Physiol 104:273–299
von Helversen D, von Helversen O (1975b) Verhaltensgenetische Untersuchungen am akustischen Kommunikationssystem der Feldheuschrecken (Orthoptera, Acrididae). II. Das Lautschema von Artbastarden zwischen *Chorthippus biguttulus* und *C. mollis*. J Comp Physiol 104:301–323
von Helversen O, von Helversen D (1987) Innate receiver mechanisms in the acoustic communication of orthopteran insects. In: Guthrie DM (ed) Aims and methods in neuroethology. Manchester Univ Press, Manchester, pp 104–150
von Helversen O, von Helversen D (1994) Forces driving coevolution of song and song recognition in grasshoppers. In: Schildberger K, Elsner N (eds) Neural basis of behavioural adaptations. G. Fischer, Stuttgart, pp 253–284
von Helversen D, von Helversen O (1995) Acoustic pattern recognition and orientation in orthopteran insects: parallel or serial processing. J Comp Physiol A 177: 767–774
von Helversen D, von Helversen O (1997) Recognition of sex in the acoustic communication of the grasshopper *Chorthippus biguttulus* (Orthoptera, Acrididae). J Comp Physiol A 180:373–386
von Helversen D, von Helversen O (1998) Acoustic pattern recognition in a grasshopper: processing in the frequency or time domain? Biol Cybern 79:467–476
Webb B, Wessnitzer J, Bush SL, Schul J, Buchli J, Ijspeert A (2007) Resonant neurons and bushcricket behaviour. J Comp Physiol A 193:285–288
Weber T, Thorson J (1989) Phonotactic behavior of walking crickets. In: Huber F, Moore TE, Loher W (eds) Cricket behavior and neurobiology. Cornell University Press, Ithaca, pp 310–339
Wendler G (1989) Acoustic orientation in crickets in the presence of two sound sources. Naturwissenschaften 76:128–129
Wiley RH (2006) Signal detection and animal communication. Adv Study Behav 36:217–247
Windmill JFG, Göpfert MC, Robert D (2005) Tympanal travelling waves in migratory locusts. J Exp Biol 208: 157–168
Wohlgemuth S, Ronacher B (2007) Auditory discrimination of amplitude modulations based on metric distances of spike trains. J Neurophysiol 97:3082–3092
Wohlgemuth S, Vogel A, Ronacher B (2011) Encoding of amplitude modulations by auditory neurons of the

locust: influence of modulation frequency, rise time, and modulation depth. J Comp Physiol A 197:61–74
Yost WA (2000) Fundamentals of hearing – an introduction. Academic, San Diego/New York
Zorović M, Hedwig B (2011) Processing of species-specific auditory patterns in the cricket brain by ascending, local, and descending neurons during standing and walking. J Neurophysiol 105:2181–2194

Auditory Prosthesis

Johan H. M. Frijns and Jeroen J. Briaire
ENT Department, Leiden University Medical Center, Leiden, The Netherlands

Synonyms

Bionic ear; Cochlear implant; Inner ear prosthesis

Definition

An auditory prosthesis is an implantable device used to (partially) restore the auditory function in people with a severe to profound hearing loss by electrically stimulating the auditory neural pathway. The cochlear implant, stimulating the auditory nerve from within the cochlea, is widely accepted as the standard rehabilitation device for this population. An auditory brain stem implant uses the same technology to stimulate the neurons of the cochlear nucleus in the brain stem and is used when the cochlea is not accessible (e.g., due to ossification after meningitis or severe hypoplasia) or the cause of deafness is found in the internal auditory canal (bilateral acoustic neuroma, aplasia of the auditory nerve).

Detailed Description

Background

Due to their reduced oral communication skills, severe to profoundly deaf people are restricted in their social functioning. Since the pioneering work of Djourno and Eyries (1957) and House in the 1970s of the last century (House and Urban 1973), it has become possible to restore some of the hearing functions through direct electrical stimulation of the auditory nerve. Currently used cochlear implants utilize electrode arrays with 12–22 contacts on a Silastic carrier that are most commonly inserted into the scala tympani through either the round window membrane or a drilled cochleostomy in its vicinity. This allows taking advantage of the tonotopic organization of the cochlea and the auditory nerve, where the high frequencies are encoded at the basal end, while the low frequencies are encoded more to the apical end (Ruggero 2009). In this way, each electrode contact of a multichannel implant aims to stimulate a different neural population, which physiologically encodes a certain pitch as determined by its intracochlear position. With current devices, however, the wish to encode all spectral information relevant for speech understanding leads to a mismatch between their tonotopic map and the physiological one.

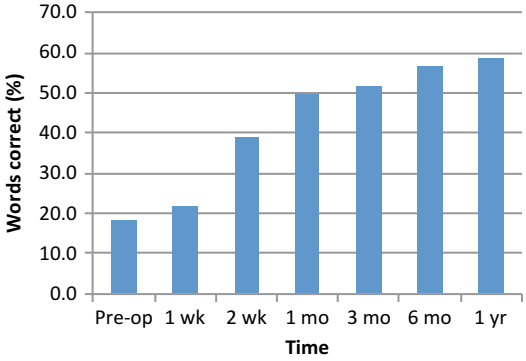
In 1984, the first cochlear implant obtained FDA approval for implantation in adults. This was followed by an NIH consensus in 1995 stating that cochlear implantation is a proven and effective rehabilitation method for deaf children and deaf adults. Today, over 188,000 people have received a cochlear implant (http://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=83). Initially, the implants provided a signal function and an aid in lipreading. Nowadays, driven by improved electronics and speech coding strategies, better electrodes and changes in inclusion criteria, the majority of the recipients achieves open set speech understanding and is able to use the telephone, although this still requires an intensive rehabilitation process (see Fig. 1).

Components and Signal Processing of Multichannel Implants

A cochlear implant consists of an external part and an internal part, as shown in Fig. 2.

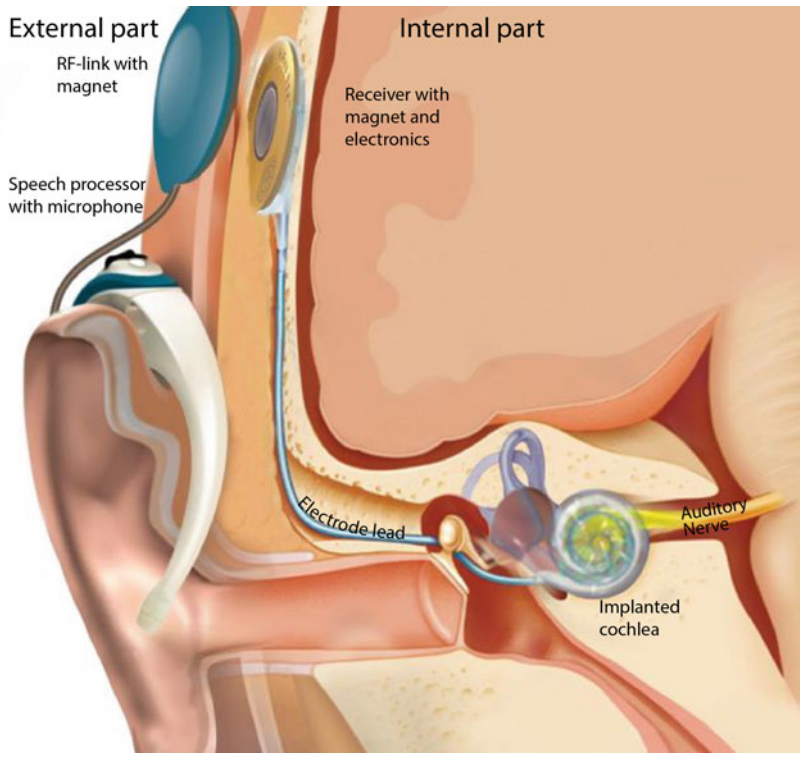
An otolaryngologist surgically implants the internal part (the so-called receiver-stimulator package with the electrode array) under general

anesthesia; the externally worn speech processor (body worn or behind the ear) is connected after several weeks of wound healing. The speech processor captures the incoming sound and, after preprocessing (typically noise cancellation and amplitude compression), encodes it into frequency-specific electrical information to be sent to the individual electrode contacts in the cochlea.



Auditory Prosthesis, Fig. 1 Performance over the first year of cochlear implant use of 70 consecutive patients implanted with a HiRes90K implant with a HiFocus 1 J electrode array (Advanced Bionics, Valencia, CA). The bars represent the percentage of correctly understood Dutch monosyllabic (CVC) words, presented from CD (65 dB SPL, free-field in quiet). The preoperative data were obtained with the best-fitted hearing aid

The speech coding strategies used in all current implants are extensions of the continuous interleaved sampling (CIS) strategy (Wilson et al. 1991). This strategy tries to avoid electrical interaction between neighboring electrode contacts by stimulating all contacts in a sequential mode rather than simultaneously. A digital filter bank is used to process the signal into separate frequency bands. Next, the envelope of each band, determined by rectification and low-pass filtering of the signal, is used to set the amplitude of a sequence of nonsimultaneous pulses on the implanted electrode contacts. The rate of stimulation is determined by the device brand and by the patient's performance, but typically ranges between 400 pulses/s and 4000 pulses/s per channel.



Auditory Prosthesis, Fig. 2 The components of a cochlear implant with a behind-the-ear speech processor. (Courtesy of Advanced Bionics)

Both the encoded stimulation pattern and the energy are transmitted to the implanted receiver-stimulator package through an RF-link. The external and internal coils for this RF-link are kept aligned with magnets in the center of both coils. The signal is picked up by the electronics in the receiver-stimulator package, which in turn delivers electrical pulses to the auditory nerve fibers via electrode contacts in the electrode array.

Modern cochlear implants also have back telemetry, allowing to record electrically evoked compound action potentials (eCAPs) of the auditory nerve via the implanted electrode array.

Computational Modeling

To provide more insight in the fundamentals of functional electrical stimulation of the auditory nerve, computational models have been developed. This involves stimulating not only the response of a nerve fiber to an externally applied potential field but also the calculation of this potential distribution from the currents on the stimulating electrodes. This is especially intricate in the case of cochlear implants due to the complex geometry of the inner ear.

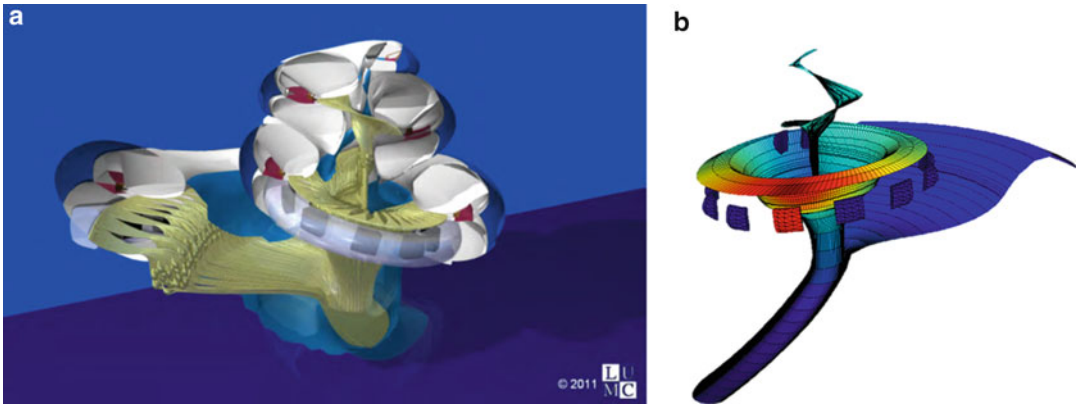
Electrical Volume Conduction in the Cochlea

An analytic solution of such a 3D volume conduction problem is restricted to geometries that are much simpler than the cochlea, and many theoretical models on the (actually three-dimensional) potential pattern set up in the cochlea by the stimulating current sources assumed an exponential decay of current from its source to the nerve fibers along the cochlea, modeled in one dimension (O'Leary et al. 1985), while other analytical approaches assume a simplified unrolled anatomy (Goldwyn et al. 2010). Suessermann and Spelman (1993) used an electrical network as a practical representation of the electro-anatomy of the cochlea.

Numerical methods nowadays, however, allow to incorporate much more detailed (electro-) anatomical information (including the shape and position of the electrode array) and sometimes even allow for patient-specific modeling on the basis of CT-scans (Carlyon et al. 2010). The numerical methods that have been used include the finite difference method (Whiten 2007), the finite element method (Rattay et al. 2001; Hanekom 2001), and the boundary element method, also known as the integral equation method (Frijns et al. 2001).

Simulating the Auditory Nerve Fiber Responses

Colombo and Parkins (1987) were the first to develop a cable model of the mammalian auditory nerve neuron based on the classical work on amphibian nerve fibers of Frankenhæuser and Huxley (1964). In order to fine tune the model to represent physiological data obtained from single auditory nerve fiber experiments in squirrel monkeys, they had to adapt the modeled nerve fiber's anatomy significantly. Motz and Rattay (1986) used a single-node model with the Hodgkin and Huxley model of unmyelinated squid giant axon membrane (▶ "Hodgkin-Huxley Model") to investigate the time structure of the response of the (myelinated!) auditory nerve to electrical stimuli. The gSEF model (Frijns et al. 1995) is a nonlinear cable model, which represents essential mammalian nerve fiber properties, including spike duration and conduction velocity, refractory behavior, and repetitive firing, better than previous models and can deal with arbitrary stimulus wave forms. It is based upon voltage clamp measurements in rat and cat motor nerve fibers at mammalian body temperature performed by Schwarz and Eikhof (1987). The gSEF model and its variants have, in conjunction with electrical volume conduction models, been used not only to predict which (intact or degenerated) fibers are excited by specific patterns of electrical stimulation (Smit et al. 2010; Frijns et al. 2009a, 2011) or to explain the results obtained with psychophysical experiments (Carlyon et al. 2010; Snel-Bongers et al. 2013) but also to calculate the eCAP produced on the basis of predicted single fiber action potentials (Briaire and Frijns 2005; Westen et al. 2011).



Auditory Prosthesis, Fig. 3 (**a**) The structure of a 3D volume conduction model of the implanted human cochlea (as developed at the Leiden University Medical Center), including the auditory nerve (in *yellow*) and a realistic representation of the electrode array in the scala tympani. (**b**) The potential distribution in the neural compartment due to monopolar stimulation

The abovementioned neural models have in common that they are deterministic in the way they treat the neural membrane responses. If the focus of research is more on the effect of high stimulation rates or on repetitive near-threshold stimulation, stochastic models come into play. Most models of this type are single-node threshold models (Bruce et al. 1999), while cable models (Rubinstein et al. 1999; Imennov and Rubinstein 2009), although computationally very intensive and requiring supercomputers, can give insight in more complex stimulation patterns.

Integrated Use of Volume Conduction and Neural Models: State of the Art

While in the early days of cochlear implantation all insights in the mechanisms underlying their function had to come either from clinical practice and associated psychophysics or from animal experiments, nowadays sophisticated computational models exist, which integrate a model of electrical volume conduction in the cochlea with active neural models. Such models can not only be used to explain effects of current and future electrode designs and stimulation schemes but are also able to predict the of anatomical variations (Frijns et al. 2009b), species differences (Frijns et al. 2001), and the effects of neural degeneration (Briaire and Frijns 2006; Snel-Bongers et al. 2013) (Fig. 3).

Cross-References

▶ Hodgkin-Huxley Model

References

Briaire JJ, Frijns JHM (2005) Unravelling the electrically evoked compound action potential. Hear Res 205(1–2):143–156

Briaire JJ, Frijns JHM (2006) The consequences of neural degeneration regarding optimal cochlear implant position in scala tympani: a model approach. Hear Res 214(1–2):17–27

Bruce IC, Irlicht LS, White MW, O'Leary SJ, Dynes S, Javel E, Clark GM (1999) A stochastic model of the electrically stimulated auditory nerve: pulse-train response. IEEE Trans Biomed Eng 46(6):630–637

Carlyon RP, Macherey O, Frijns JHM, Axon PR, Kalkman RK, Boyle P, Baguley DM, Briggs J, Deeks JM, Briaire JJ, Barreau X, Dauman R (2010) Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear. J Assoc Res Otolaryngol 11(4):625–640

Colombo J, Parkins JW (1987) A model of electrical excitation of the mammalian auditory-nerve neuron. Hear Res 31:287–312

Djourno A, Eyries C (1957) Auditory prosthesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelt coiling. Presse Med 65:1417

Frankenhæuser B, Huxley AF (1964) The action potential in the myelinated nerve fiber of *Xenopus laevis* as computed on the basis of voltage clamp data. J Physiol Lond 171:302–315

Frijns JHM, de Snoo SL, Schoonhoven R (1995) Potential distributions and neural excitation patterns in a rotationally symmetric model of the electrically stimulated cochlea. Hear Res 87:170–186

Frijns JHM, Briaire JJ, Grote JJ (2001) The importance of human cochlear anatomy for the results with modiolus hugging multi-channel cochlear implants. Otol Neurotol 22(3):340–349

Frijns JHM, Kalkman RK, Vanpoucke FJ, Bongers JS, Briaire JJ (2009a) Simultaneous and non-simultaneous dual electrode stimulation in cochlear implants: evidence for two neural response modalities. Acta Otolaryngol 129(4):433–439

Frijns JH, Kalkman RK, Briaire JJ (2009b) Stimulation of the facial nerve by intracochlear electrodes in otosclerosis: a computer modeling study. Otol Neurotol 30(8): 1168–1174

Frijns JHM, Dekker DMT, Briaire JJ (2011) Neural excitation patterns induced by phased-array stimulation in the implanted human cochlea. Acta Otolaryngol 131(4):362–370

Goldwyn JH, Bierer SM, Bierer JA (2010) Modeling the electrode-neuron interface of cochlear implants: effects of neural survival, electrode placement, and the partial tripolar configuration. Hear Res 268(1–2):93–104

Hanekom T (2001) Three-dimensional spiraling finite element model of the electrically stimulated cochlea. Ear Hear 22(4):300–315

House WF, Urban J (1973) Long term results of electrode implantation and electronic stimulation of the cochlea in man. Ann Otol Rhinol Laryngol 82(4): 504–517

Imennov NS, Rubinstein JT (2009) Stochastic population model for electrical stimulation of the auditory nerve. IEEE Trans Biomed Eng 56(10):2493–2501

Motz H, Rattay F (1986) A study of the application of the Hodgkin-Huxley and the Frankenhaeuser-Huxley model for electrostimulation of the acoustic nerve. Neuroscience 18:699–712

O'Leary SJ, Black RC, Clark GM (1985) Current distributions in the cat cochlea: a modelling and electrophysiological study. Hear Res 18:273–281

Rattay F, Leao RN, Felix H (2001) A model of the electrically excited human cochlear neuron. II. Influence of the three-dimensional cochlear structure on neural excitability. Hear Res 153(1–2):64–79

Rubinstein JT, Wilson BS, Finley CC, Abbas PJ (1999) Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. Hear Res 127(1–2):108–118

Ruggero MA (2009) Cochlea. In: Binder MD, Hirokawa N, Windhorst U (eds) Encyclopedia of neuroscience. Springer, Berlin Heidelberg, pp 765–769

Schwarz JR, Eikhof G (1987) Na currents and action potentials in rat myelinated nerve fibres at 20 °C and 37 °C. Pflugers Arch 409:569–577

Smit JE, Hanekom T, van Wieringen A, Wouters J, Hanekom JJ (2010) Threshold predictions of different pulse shapes using a human auditory nerve fibre model containing persistent sodium and slow potassium currents. Hear Res 269(1–2):12–22

Snel-Bongers J, Briaire JJ, Van Veen EH, Kalkman RK, Frijns JHM (2013) Threshold levels of dual electrode stimulation in cochlear implants. J Assoc Res Otolaryngol 14:781–790

Suesserman MF, Spelman FA (1993) Lumped-parameter model for in vivo cochlear stimulation. IEEE Trans Biomed Eng 40:237–245

Westen AA, Dekker DMT, Briaire JJ, Frijns JHM (2011) Stimulus level effects on neural excitation and eCAP amplitude. Hear Res 280:166–176

Whiten D (2007) Electro-anatomical models of the cochlear implant. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA

Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991) Better speech recognition with cochlear implants. Nature 352(6332): 236–238

Further Reading

NIH factsheet cochlear implants: http://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=83

Auditory Recognition

▶ Acoustic Timbre Recognition

Auditory Scene Analysis

▶ Auditory Cortex: Separating Signal from Noise
▶ Auditory Perceptual Organization

Auditory Scene Segmentation

▶ Auditory Cortex: Separating Signal from Noise

Auditory Sensory Receptor Cell, Model

► Cochlear Inner Hair Cell, Model

Auditory Thalamocortical Transformations

Kazuo Imaizumi and Charles C. Lee
Department of Comparative Biomedical Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA, USA

Definition

Auditory thalamocortical transformations arise from the ascending neural processing of spontaneous activity as well as external sound-evoked activity from the auditory thalamus, the medial geniculate body, to the thalamorecipient layers of the auditory cortex in mammals.

Detailed Description

Thalamic and Cortical Organization

The Thalamus

The thalamus is the obligate neural station conveying ascending sensory information to the cortex (Sherman and Guillery 2006; Jones 2007). Each sensory modality, except for olfaction, is represented in defined nuclei of the thalamus through which sensory information must first be processed before eventually being transmitted to the respective sensory areas of the neocortex (Sherman and Guillery 2006; Jones 2007). In the auditory system, the medial geniculate body is the main auditory thalamic nucleus (Winer 1984; Jones 2007). Classically, the medial geniculate body has been divided into three main subdivisions, i.e., the ventral division, the dorsal division, and the medial division (Winer 1984; Imig and Morel 1985). Each of these thalamic nuclei can be identified on the basis of their cytoarchitecture, physiological properties, and connections (Huang and Winer 2000; de la Mothe et al. 2006; Lee and Winer 2008a). The thalamocortical transformation is constrained by the neuroanatomical organization of projections from each of these thalamic nuclei to each of the several areas of the auditory cortex (Winer et al. 2005; Winer and Lee 2007; Lee and Winer 2011b).

The Ventral Division of the Medial Geniculate Body

The ventral division of the medial geniculate body is the principal thalamic nucleus conveying ascending auditory information to the primary auditory cortex (Huang and Winer 2000; Smith et al. 2012). A subregion within the ventral division, the pars ovoidalis, is located in the medial part of the nucleus, bordering the dorsal division and medial division (Winer 1984; Jones 2007). Neurons in the ventral division are sharply tuned to sound frequencies (Imig and Morel 1985). These neurons are arranged in laminar rostrocaudal sheets, with their dendritic fields aligned in parallel along the sheet. The neurons in each sheet respond to similarly tuned frequencies (Imig and Morel 1985). These sheets are organized lateromedially in most species studied, with neurons in the lateral regions responding to lower frequencies of sound, while neurons in the medial regions respond to higher frequencies of sound (Imig and Morel 1985). Other physiological properties, such as bandwidth tuning and aurality, are found interdigitated among neurons across these isofrequency laminae (Ehret 1997). The main ascending projection to the ventral division originates from the central nucleus of the inferior colliculus (Calford 1983; Wenstrup 2005). These projections preserve the topographic organization connecting similarly tuned regions in the inferior colliculus to matched regions in the ventral division of the medial geniculate body (Wenstrup 2005). The main output of the ventral division is to the primary auditory cortex. These projections terminate primarily in layer 4 of the

primary auditory cortex but also have branched projections to layer 6 (Huang and Winer 2000; Smith et al. 2012). Similar to the projection from the inferior colliculus, these projections are also topographically organized, such that the low-frequency regions of the ventral division project to the lower-frequency regions of the primary auditory cortex (Morel and Imig 1987; Morel et al. 1993). The ventral division also connects with other tonotopically organized auditory cortical areas, which vary in number in different species, e.g., five in the cat (Reale and Imig 1980) and three in the monkey (Kaas and Hackett 2000). These areas also receive topographically organized projections from the ventral division but mainly from nonoverlapping sectors (Morel and Imig 1987). Very few neurons in the ventral division send branched projections to multiple auditory areas, i.e., low-frequency neurons in the ventral division do not connect with multiple low-frequency regions in different tonotopically organized areas (Lee et al. 2004, 2011).

The Dorsal Division of the Medial Geniculate Body

The dorsal division of the medial geniculate body is the part of the auditory thalamus that connects to non-tonotopically organized areas of the auditory cortex (de la Mothe et al. 2006; Lee and Winer 2008a). This nucleus is composed of various subnuclei, including the dorsal superficial, deep dorsal, dorsocaudal, and ventrolateral nuclei (Winer et al. 2005; Lee and Winer 2008a). Neurons in the dorsal division are broadly tuned to frequencies, many with multi-peaked and complex receptive fields (Morel and Imig 1987; Winer et al. 2005). In contrast to the organization of the ventral division of the medial geniculate body, neurons in the dorsal division do not exhibit an oriented laminar pattern of organization (Winer 1984). The dendritic arborizations of neurons in the dorsal division are more isotropically organized. The subdivision of this nucleus is based primarily on cytoarchitectonic densities and connections with several non-tonotopic auditory areas (Lee 2013). Projections to the dorsal division originate primarily from the non-tonotopic area of the inferior colliculus, i.e., the dorsal cortex (Wenstrup 2005). The dorsal division nuclei project broadly to the non-tonotopic areas of the auditory cortex, which also include multimodal and limbic-related areas, whose relation to auditory processing in part is derived from the direct inputs received from thalamocortical sources (Lee and Winer 2008a, 2011a). Although a functional metric, such as tonotopy, appears to be absent in the nuclei of the dorsal division, their projections to the non-tonotopic areas of the auditory cortex are highly topographic, similar in extent to the topography of projections from the ventral division to the tonotopic regions of the auditory cortex (Lee and Winer 2005; Schreiner and Winer 2007). The termination pattern of the dorsal division projections to the non-tonotopic cortical areas, in particular, the secondary auditory cortical region, is similar to that of the ventral division projection to the primary auditory cortex, i.e., terminations primarily in layers 4 and 6 (Huang and Winer 2000). However, the synaptic terminals of the dorsal division projection to the secondary auditory cortex are slightly larger on average than the projections from the ventral division to the primary auditory cortex (Smith et al. 2012). Again, although the dorsal division projects broadly to several non-tonotopic auditory areas, very few neurons within the dorsal division send branched projections to multiple areas (Kishan et al. 2008; Lee and Winer 2008a; Lee et al. 2011).

The Medial Division of the Medial Geniculate Body

The medial division of the medial geniculate body projects widely across all auditory cortical areas, with a pattern of axonal termination that contrasts with the projections from the ventral division and the dorsal division of the medial geniculate body (Huang and Winer 2000; Jones 2007). Neurons in the medial division exhibit highly complex receptive fields, which often extend beyond purely auditory responses, with many neurons responding to stimuli from several modalities, i.e., visual and somatosensory (Bordi and LeDoux 1994). Neurons in the medial division differ from those in the ventral division and dorsal division, in that they display a wide range of sizes and dendritic arborization patterns (Bartlett and Smith 1999; Smith et al. 2007). The medial division contains the magnocellular neurons, which are

the largest neurons in the medial geniculate body (Winer 1984). Like the dorsal division, the neurons in the medial division do not appear to be organized isotropically along any particular anatomical domain and are loosely packed compared with the other divisions (Winer 1984). The medial division receives its primary input from both the dorsal cortex and external cortex regions of the inferior colliculus, which also contain nontonotopic and multimodal responsive neurons (Wenstrup 2005). Every area of the auditory cortex receives a projection from the medial division (Lee and Winer 2008a), but unlike the ventral division and the dorsal division, the laminar terminations of axons in these areas are primarily concentrated in layer 1 (Huang and Winer 2000). While axonal divergence from the medial geniculate body is low in general, the highest proportion of neurons projecting to multiple auditory cortical areas is found in the medial division; however, these comprise on average less than 2% of neurons in the medial division of the medial geniculate body (Kishan et al. 2008, 2011; Lee et al. 2011).

Inhibitory Circuits in the Thalamus

The thalamic reticular nucleus, although not a specific constituent of the medial geniculate body, is intimately intertwined with the operations of both the thalamus and cortex and thus is an essential structural component of the thalamocortical transformation (Winer and Larue 1996; Crabtree et al. 1998; Pinault 2004; Sherman and Guillery 2006). The thalamic reticular nucleus is composed of inhibitory GABAergic neurons that form a shell surrounding the thalamus, roughly located along the lateral border of the thalamus and extending rostrocaudally along nearly its entire length (Pinault 2004; Lam and Sherman 2005, 2007, 2010). In general, each nucleus of the thalamus innervates a specific sector of the thalamic reticular nucleus and receives reciprocal topographic inhibitory feedback projections from that region of the thalamic reticular nucleus (Lam and Sherman 2005). In addition, feedback projections from layer 6 of the neocortex en route to the thalamus branch to innervate the thalamic reticular nucleus (Lam and Sherman 2010), which establishes an extended thalamocorticothalamic inhibitory feedback loop (Pinault 2004). The thalamic reticular nucleus and the inferior colliculus are the main sources of inhibition in the medial geniculate body of rodents, which contain very few local inhibitory neurons (Winer and Larue 1996). In humans, local inhibitory neurons in the medial geniculate body comprise ~20% of the total number of neurons, which is accompanied by a proportional reduction in inhibitory projections from the thalamic reticular nucleus, compared with rodents and other species, which have fewer than 1% of local inhibitory neurons in the medial geniculate body (Winer and Larue 1996).

The Cerebral Cortex

The cerebral cortex is phylogenetically the newest structure in the mammalian brain, responsible for the higher-order processing of sensory, motor, and limbic information (Kaas 2008). Broadly, the cerebral cortex is composed of regions of "gray" matter and "white" matter, the former residing near the outer cortical surface and containing all of the neuronal cell bodies and the latter residing beneath the gray matter and composed of the axonal fiber tracts of afferent and efferent projections (Nieuwenhuys 2013). Although regional variations exist, the gray matter of the cerebral cortex has a laminar organization divided into cytoarchitectonically distinguishable layers, such that neuronal cell bodies are situated in structural and functional groups relative to their location along the pial to white matter axis (Mountcastle 1997). These layers of neuronal cell bodies have specific afferent and efferent connections with other cortical regions and with subcortical structures, in particular the thalamus (Sherman and Guillery 2006). In total, there are six classically defined layers of the cortex (Mountcastle 1997). Of these, layer 4 of the cerebral cortex is the main recipient layer for sensory information ascending from the primary sensory thalamic nuclei (Sherman and Guillery 2006). In addition, layer 6 receives branched projection from these primary sensory thalamic nuclei (Huang and Winer 2000; Smith et al. 2012; Lee and Imaizumi 2013), and layer 1 is the recipient of thalamocortical inputs

from nonspecific thalamic nuclei, e.g., the medial division of the medial geniculate body (Huang and Winer 2000; Jones 2007). In addition, layer 6 is the source of neurons that send feedback projections to the thalamic nucleus that provides its main thalamocortical input in layers 4 and 6, which establishes a thalamocorticothalamic feedback loop (Sherman and Guillery 2006). Cortical layer 5 is the source of feedforward nonreciprocal thalamocortical projections, which serve as a conduit for communication between cortical areas via a corticothalamocortical route (Sherman and Guillery 2006).

The surface of the cerebral cortex is regionally specified into distinct functional areas that are involved in processing sensory and motor information (Kaas 2008). The boundaries of these cortical areas, including those involved with audition, are broadly defined based on their cytoarchitectural organization, connections with other structures, and physiological responses of constituent neurons, although precise borders and definitions for many cortical areas in different organisms still remain elusive (Kaas and Hackett 2000; Hackett 2011). Although all mammals have cortical regions devoted to the processing of auditory information, the number of auditory cortical areas varies widely among different species (Lee and Winer 2008b, 2011b). Despite this heterogeneity, all mammals studied thus far have an identifiable auditory cortical region that receives direct input from the ventral division of the medial geniculate body (Lee and Winer 2011b). This primary auditory cortical area is also defined on the basis of a clearly identifiable tonotopic map, which reflects the frequency segregation established in the cochlea and propagated along the auditory pathway (Ehret 1997).

Although the primary auditory cortical area is the most conserved across species, constellations of other cortical regions devoted to the processing of sound exist among the cortices of different mammalian species (Kaas 2008). However, their physiological and anatomical organization is generally less well understood relative to that of the primary auditory cortex. Nevertheless, these other areas can be broadly grouped into distinct categories: tonotopic, non-tonotopic, multimodal, and limbic related (Lee and Winer 2011a). The tonotopic regions, like the primary auditory cortex, contain identifiable maps of frequency across their surface, generally with frequency reversals at their borders, and receive strong projections from the ventral division of the medial geniculate body (Reale and Imig 1980; Hackett et al. 1998; Hackett et al. 2011). The non-tonotopic areas contain disordered representation of frequency and generally receive more prominent inputs from the dorsal division of the medial geniculate body (Schreiner and Cynader 1984; Smith et al. 2012). Multimodal and limbic areas reside at the limits of the classical auditory areas and receive inputs from multimodal and limbic thalamic nuclei and areas and have complex responses reflective of these convergent inputs (Bowman and Olson 1988; Clarey and Irvine 1990; Clascá et al. 1997; de la Mothe et al. 2006). As discussed above, features of the thalamocortical inputs to these other areas resemble those to the primary auditory cortex (Huang and Winer 2000; Smith et al. 2012), which could serve as an anatomical basis for common thalamocortical transformations across the expanse of auditory cortical areas (Lee and Sherman 2008, 2011).

Receptive Fields and the Thalamocortical Transformation

Spectral Receptive Field

An important physiological parameter in the auditory thalamocortical transformation is the spectral receptive field. The spectral receptive field is, in general, measured based on a frequency-threshold tuning curve (response to sound level as a function of sound frequency). A common measure is the Q factor by which characteristic frequency is divided by a linear measure of bandwidth at a given sound level above threshold (e.g., $Q10$; Q value at 10 dB above threshold) (Imaizumi et al. 2004). Because the Q value is a normalized measure, the larger the Q value, the more sharply tuned are the neurons. Another measure is the square root transformation:

$$\sqrt{F2} - \sqrt{F1}$$

where $F2$ and $F1$ are the highest and lowest edges of bandwidth at a given sound level (Rouiller et al. 1981; Calford 1983). Unlike the Q value, the smaller the square root transformation value, the more sharply tuned are the neurons.

The auditory thalamocortical transformation is mediated by the excitatory neurotransmitter, glutamate, from the medial geniculate body to the auditory cortex. Therefore, only the excitatory receptive field is transformed in thalamocortical projections. A problem describing spectral receptive fields in the thalamocortical transformation is the availability of data utilizing the same recording method (e.g., local field potentials, single- or multiunit extracellular recording, whole-cell recording, etc.) under similar recording conditions (e.g., anesthesia: isoflurane, ketamine, or pentobarbital; depth of anesthesia; awake states; restricted or freely moving) in the same animal species. This creates some controversy and readers should use caution for the following section.

Functional organization of the spectral receptive field has been well appreciated in the auditory cortices of several species. The most well-known example is the primary auditory cortex of the mustached bat. A large area, called the DSCF (Doppler-shifted constant frequency) area, of the primary auditory cortex is devoted to a particular frequency range (60.6–62.3 kHz) for their prey-hunting behavior (Suga 1994). Neurons in the DSCF area are extremely sharply tuned to characteristic frequency ($Q50$ values range from ~10 to 500 or higher) (Suga and Manabe 1982; Suga et al. 1997). Neurons in the anterior and posterior parts of the DSCF area are more broadly tuned. Similarly organized clusters of sharply or broadly tuned neurons are also found in the auditory cortex of carnivores and primates (Recanzone et al. 1999; Cheung et al. 2001; Read et al. 2001; Imaizumi et al. 2004; Philibert et al. 2005; Imaizumi and Schreiner 2007). In particular, the cat primary auditory cortex has an interesting functional organization of spectral receptive fields. Sharply or broadly tuned neurons are clustered alternatively along the dorsoventral axis only in the mid-frequency range (5–20 kHz) (Imaizumi and Schreiner 2007). Unlike those animal species, functional organization of spectral receptive fields in the rodent auditory cortex is not clear (Polley et al. 2007). Whereas rich information of functional organization of spectral receptive fields is available in the auditory cortex, only limited information is available in the medial geniculate body. Neurons in the ventral division are, in general, more sharply tuned than in the dorsal division (Rouiller et al. 1981; Calford 1983; Edeline et al. 1999). However, no clear spatial organization is available due to the deep location in the brain.

In general, it is believed that spectral receptive fields become broader through the thalamocortical transformation (Kaur et al. 2004, 2005). However, more recent studies using in vivo whole-cell recordings combined with pharmacological manipulation (blocking cortical input by muscimol and SCH50911) have revealed that thalamocortical input is not as sharp as was thought (Liu et al. 2007). These examples above are based on studies in rodents. Spectral receptive fields, however, vary widely across species. In the human auditory cortex, neurons are extremely sharply tuned (Bitterman et al. 2008), which suggests that spectral receptive fields may become sharper through the thalamocortical transformation. How does this opposite trend occur in thalamocortical transformation? Clear evidence is available in the awake mustached bat. Based on $Q10$, $Q30$, and $Q50$ values, DSCF neurons in the primary auditory cortex are more sharply tuned than those in the medial geniculate body (Suga et al. 1997). Similar evidence is also available in the awake marmoset (Bartlett et al. 2011). Thus, the sharpening of spectral receptive fields through the thalamocortical transformation may relate to behavioral and/or ethological functions.

For thalamocortical transformation of spectral receptive fields, limited experimental cases are available using in vivo single-unit recordings in awake guinea pigs (Creutzfeldt et al. 1980) and anesthetized cats (Miller et al. 2001, 2002). For secure functional transformation, medial geniculate body and auditory cortex neurons require an alignment of less than 1/3 octave difference in best frequency (sound frequency evoked the best response in a neuron at a given sound level)

(Miller et al. 2001). To fully activate a neuron in the auditory cortex, synaptic convergence from 20 to 25 neurons in the medial geniculate body is required (Miller et al. 2001). Using a ripple noise stimulus and computational analytical approach of reverse correlation technique, Miller et al. (2002) estimated spectral *modulation* rates through thalamocortical transformation. Both thalamic and cortical neurons show lower spectral *modulation* rates. Whereas cortical neurons have significantly lower spectral *modulation* rates than thalamic neurons based on the best spectral *modulation* rates, the overall spectral filter properties between thalamic and cortical neurons are similar.

Overall, spectral receptive (*modulation*) fields in the thalamocortical transformation are not simple. Depending on animals and the behavioral significance of sound frequency, spectral receptive fields become broader or sharper through the thalamocortical transformation.

Temporal Receptive Field

An important function of the central auditory system is to decode species-specific communications and human speech sounds. Periodic modulations are ubiquitous temporal features of species-specific communications and human speech sounds (Rosen 1992; Joris et al. 2004). Measuring repetition-rate transfer functions or modulation transfer functions captures response characterization to assess information for a temporal range that corresponds to periodicities in communication sounds (Eggermont 2001; Joris et al. 2004). Two commonly used measures to characterize repetition-rate transfer functions are firing rate and vector strength (VS). Firing rate estimates overall response magnitude in a particular time window. Vector strength estimates spike-timing precision to a particular phase of repetition or modulation stimulus:

$$VS = \frac{\sqrt{(\sum \cos\theta)^2 + (\sum \sin\theta)^2}}{n}$$

$$\theta = 2\pi \frac{t}{T}$$

where n is the total number of spikes, t is time of spike occurrence, and T is the interstimulus interval (Goldberg and Brown 1969). Significance of synchronization to the stimulus is examined by a Rayleigh test, >13.8 ($p < 0.001$) (Mardia 1972). Because vector strength does not incorporate response strength, it may give rise to high vector strength values when the response strength is low. To overcome a shortcoming of the VS measure, phase-projected vector strength (VS_{pp}) is used (Yin et al. 2011; Niwa et al. 2012). VS_{pp} compares the mean phase angle for each trial with the mean phase angle of all trials and penalizes single-trial vector strength values if they are not in phase with the global response. VS_{pp} is computed on a trial-by-trial basis as follows:

$$VS_{pp} = VS_t \cos(\phi_t - \phi_c)$$

where VS_{pp} is the phase-projected vector strength per trial, VS_t is the vector strength per trial, and ϕ_t and ϕ_c are the trial-by-trial and mean phase angle in radians, calculated for each stimulus condition, and.

$$\phi = arctan\, 2 \frac{\sum_{i=1}^{n} \sin\theta i}{\sum_{i=1}^{n} \cos\theta i}$$

where n is the number of spikes per trial (for ϕ_t) or across all trials (for ϕ_c) and arctan2 is a modified version of the arctangent that determines the correct quadrant of the output based on the signs of the sine and cosine inputs. Whereas vector strength value ranges from 1 (all spikes occur at the same stimulus phase) to 0 (spikes occur randomly), phase-projected vector strength value ranges from 1 (all spikes in phase with population mean phase) to -1 (all spikes 180° out of phase with population mean phase) with 0 corresponding to randomly occurring spikes (Yin et al. 2011; Niwa et al. 2012). Another way to overcome a shortcoming of vector strength is the product of two measures (vector strength and firing rate), i.e., phase-locked rate or synchronized rate (Eggermont 1998b; Joris et al. 2004; Imaizumi et al. 2011). The synchronized rate measure incorporates both timing and response strength measures.

Two coding schemes for temporal receptive fields have been proposed: precise spike timing estimated by vector strength codes slow repetition rates (or modulation frequencies), while firing rate codes faster repetition rates (De Ribaupierre et al. 1972; Bieser and Muller-Preuss 1996; Schulze and Langner 1997; Lu and Wang 2000; Lu et al. 2001; Joris et al. 2004). However, more recent studies have proposed different coding schemes, as will be described later.

A problem describing temporal receptive fields in the thalamocortical transformation is also the availability of data utilizing the same stimulus (e.g., nature of stimulus: click or white noise; modulation carrier: pure tone or noise; modulation depth; and modulation shape: rectangular or sinusoidal), stimulus presentation method (free field or sealed earphones), and recording method (single- or multiunit extracellular recording) under similar recording conditions (e.g., anesthesia: isoflurane, ketamine, or pentobarbital; depth of anesthesia; awake states: restricted, freely moving, or engaged behavior) in the same animal species. In many cases, the analytical criteria are also different. These create difficulty for direct comparisons and some discrepancy. Thus, the readers should treat the following section with caution.

The vast majority of studies regarding temporal receptive fields have focused on the primary auditory cortex. However, functional magnetic resonance imaging or positron emission tomography in humans and macaques suggested that the superior temporal plane is specific to human speech or macaque species-specific calls over nonspecific calls or other sounds (Belin et al. 2000; Poremba et al. 2004; Petkov et al. 2008). These fields are located anterior to the primary core fields and may correspond to the rostral field in primates. Reversible lesion experiments in cats also show that the anterior auditory field is specific to temporal pattern discrimination (Lomber and Malhotra 2008). Thus, the rostral field in primates and the anterior auditory field in carnivores (and rodents) may be more important for temporal pattern processing.

Rodent Auditory Cortex and Thalamus

Under anesthesia of pentobarbital, neurons in the rat primary auditory cortex, in general, show a low-pass filter property of repetition-rate transfer functions by rate coding (firing rate) or band-pass filter property by temporal coding (vector strength) (Kilgard and Merzenich 1999; Chang et al. 2005). Ter-Mikaelian et al. (2007) examined the effects of anesthesia (a combination of pentobarbital and ketamine) and awake (head-fixed) states in the gerbil primary auditory cortex. While anesthesia certainly affects modulation transfer functions in individual neurons, overall, the population response seems to be similar between anesthesia and awake states (Ter-Mikaelian et al. 2007). Clear evidence of an effect of anesthesia is available in the rat primary auditory cortex (Rennaker et al. 2007). Ketamine anesthesia significantly decreased cutoff repetition rates (higher border of repetition-rate transfer function) to $<$20 Hz from 80 to 90 Hz in the awake state. A more recent study using awake state (head-fixed) rats examined repetition-rate transfer function using click trains in the primary auditory cortex and the anterior auditory field (Ma et al. 2013). Both core fields show high best repetition rates (up to 32–64 Hz) and cutoff repetition rates (up to 256 Hz), although neurons in the anterior auditory field prefer significantly higher repetition rates than those in the primary auditory cortex. These best and cutoff repetition rates in the awake rat auditory cortex are, in general, higher than those in the anesthetized rat (Kilgard and Merzenich 1999; Chang et al. 2005). Attention or engaged behaviors also alter temporal receptive fields. By presenting at 15 Hz repetition rate (with various carrier frequencies) paired with electrical stimulation of the nucleus basalis for 20–25 days, neurons in the rat primary auditory cortex are capable of following higher repetition rates by rate coding than those in control (Kilgard and Merzenich 1998). Training in a sound maze in which rats use sound source location for food rewards based on auditory cues (noise repetition rates increased with decreasing distance between the rat and target location) also enhanced temporal receptive fields (Bao et al. 2004). Compared to

studies in the rodent auditory cortex, a small number of studies are available in rodent thalamus. In awake guinea pigs, thalamic neurons show more robust responses to higher modulation frequencies than cortical neurons (Creutzfeldt et al. 1980).

Cat Auditory Cortex and Thalamus

The cat auditory cortex has been a focus of studies of temporal pattern processing. In particular, the primary auditory cortex and the anterior auditory field are readily identifiable based on their location relative to sulcal patterns (Knight 1977; Imaizumi et al. 2004; Lee et al. 2004) and are often compared in temporal pattern processing (Schreiner and Urbas 1988; Eggermont 2000; Joris et al. 2004). Under anesthesia, neurons in the anterior auditory field prefer higher modulation frequencies (and repetition rates) than those in the primary auditory cortex (Schreiner and Urbas 1988; Eggermont 1998b). Anesthesia also reduces temporal receptive fields by half in neurons of the cat primary auditory cortex (Goldstein et al. 1959). In the cat anterior auditory field under ketamine anesthesia, Imaizumi et al. (2010) proposed different coding schemes using a combination of an in vivo high-resolution cortical mapping technique with information theory. Because this study was made using multiunit recordings, they computed discriminability of six different low repetition rates (1–30 Hz) (Imaizumi et al. 2010). Unlike the traditional coding scheme (precise spike timing codes low repetition rates), inter-spike intervals can carry much more information than timing and rate codes: some multiunits carried an information value >2 bits that is close to a maximum of ~2.58 bits ($=\log_2(6)$). Furthermore, spatial distribution of normalized firing rate to six different repetition rates differs across the stimuli, which provides a potential coding scheme in the view of an ideal observer. These results suggest concurrent coding schemes of temporal pattern processing by inter-spike intervals, firing rate, and a map (Imaizumi et al. 2010). Using behaviorally trained cats, Dong et al. (2011) compared neurometric (neural responses to six repetition rates from 12.5 to 200 Hz by in vivo single-unit recordings) with psychometric (Go/No-Go behavioral responses to the same six repetition rates) functions. Their recordings were focused on the relatively low-frequency locations <16 kHz of the primary auditory cortex. Prevalence is found of more synchronized units in the behaviorally engaged primary auditory cortex than in an anesthetized one (Lu and Wang 2000; Dong et al. 2011). However, rate coding by non-synchronized units correlates well with psychometric functions.

Compared to the studies in the cat auditory cortex, a small number of studies are available in the cat thalamus. In nitrous oxide-anesthetized cats, approximately half of the thalamic units showed precisely time-locked responses to click trains (Rouiller et al. 1981). Half of these units had cutoff repetition rates (higher border of repetition-rate transfer function) up to 100 Hz. Local field potentials in the auditory cortex record subthreshold responses potentially from thalamocortical fibers (and corticocortical fibers), thus indicating thalamic responses. Both best modulation frequencies and cutoff modulation frequencies are generally higher in local field potentials than unit recordings in the primary auditory cortex and anterior auditory field (Eggermont 1998b). Under ketamine anesthesia in the primary auditory cortex and the ventral division of the medial geniculate body of cats, Miller et al. (2002) conducted in vivo simultaneous single-unit recordings from thalamus and cortical neurons. They found that temporal modulation transfer functions are significantly deteriorated through thalamocortical projections.

Primate Auditory Cortex and Thalamus

A majority of studies of temporal receptive fields in primates are conducted in the awake state. In the awake squirrel monkey, neurons in the primary auditory cortex and the rostral field show both temporal and rate codes to amplitude modulation frequencies up to 64 or 128 Hz (Bieser and Muller-Preuss 1996). Neurons in the primary auditory cortex showed higher best modulation frequencies than those in the rostral field. In the awake macaque, neurons in the primary auditory cortex have higher best modulation frequencies by both temporal (means are 13 and 4.8 Hz) and rate (means are 45 and 19 Hz) codes and higher

vector strength than those in the rostral field (Malone et al. 2010; Scott et al. 2011). These examples from the awake squirrel monkey and macaque have suggested the opposite trend of temporal receptive fields in the core auditory fields to rodents and cats (Schreiner and Urbas 1988; Eggermont 1998b; Joris et al. 2004) despite the similar cortical locations (relative to the primary auditory cortex) of the anterior auditory field in rodents and cats and the rostral field in the primates. In the awake primate auditory cortex, neurons may carry both temporal and rate codes for lower and higher repetition rates (or modulation frequencies) (Bieser and Muller-Preuss 1996; Lu et al. 2001; Liang et al. 2002; Malone et al. 2007; Yin et al. 2011). However, a proportion of synchronized (using vector strength or similar measures) and non-synchronized (using firing rate) neurons are different among the studies. These discrepancies may be caused by the stimulus, the range of repetition rates, and/or analytical criterion (Yin et al. 2011). There is an interesting proposal of low to mid range of repetition rates (10–45 Hz) corresponding to flutter perception by two different populations of neurons in the awake marmoset auditory cortex. One population of neurons increases firing rate with increasing repetition rates, while the other population decreases firing rate with increasing repetition rates (Bendor and Wang 2007). All examples reviewed above are based on studies of primates passively listening to stimuli in awake state. However, active engagement of behaviors (discriminating modulated sounds, 2.5–500 Hz, from unmodulated sounds) improves both temporal and rate codes in single neurons of the macaque primary auditory cortex (Niwa et al. 2012).

Compared to the studies in primate auditory cortex, a small number of studies are available in primate thalamus. In the awake squirrel monkey, neurons in the thalamus show a similar tendency of temporal and rate coding to best modulation frequencies up to 128 Hz (Preuss and Muller-Preuss 1990). In the awake marmoset, neurons in the ventral and the anterodorsal divisions of the medial geniculate body show a mixture of temporal and rate coding, while neurons in the posterodorsal division show a dominant tendency of rate coding (Bartlett and Wang 2011). These examples suggest that separation of temporal and rate coding for low and high repetition rates may be created within the auditory cortex (Bartlett and Wang 2007). However, other data suggest that separation from temporal to rate coding for low to high repetition rates may be completed through thalamocortical transformation (Malone et al. 2007; Yin et al. 2011). Thalamic neurons are capable of following higher repetition rates by both temporal and rate coding than cortical neurons in the awake marmoset (Bartlett and Wang 2007).

Overall, temporal receptive fields change through the thalamocortical transformation. In general, thalamic neurons are more capable of following higher repetition rates (modulation frequencies) than cortical neurons. However, neurons in the auditory cortex may employ different coding schemes to follow different ranges of repetition rates (this is not restricted only to high repetition rates but also low to mid repetition rates) either through the thalamocortical transformation or at the cortical level.

Latency

First-spike latency (hereafter latency) is another important physiological parameter (Eggermont 2001). However, because only the experimenter knows the stimulus onset (the brain and neurons do not know it), relative latencies might be a good candidate for neural coding of temporal patterns (Eggermont 1998b; Schreiner and Raggio 1996; Lu and Wang 2000; Liang et al. 2002; Ter-Mikaelian et al. 2007; Imaizumi et al. 2011), vocalizations (Wang et al. 1995; Nagarajan et al. 2002), sound localization (Eggermont 1998a; Furukawa et al. 2000), and auditory scene (Dear et al. 1993). Latency, in general, decreases with increasing sound level (Heil 1997). However, neurons in the primary auditory cortex of the little brown bat show shorter latencies to lower sound level than higher sound level (Sullivan 1982b). Furthermore, when the two sounds (higher and lower sound levels) are presented by a behaviorally relevant delay between pulse and echo, latencies are facilitated and become shorter for the

echolocating behavior (Sullivan 1982a). In general, neurons in the anterior auditory field have shorter latencies than those in the primary auditory cortex across many different species (Schreiner and Urbas 1986, 1988; Linden et al. 2003; Rutkowski et al. 2003; Imaizumi et al. 2004; Bizley et al. 2005). However, neurons in the primate primary auditory cortex show shorter latencies than those in the rostral field (Scott et al. 2011), which is related to the fact that neurons in the primate primary auditory cortex follow higher repetition rates than those in the rostral field (Bieser and Muller-Preuss 1996; Malone et al. 2010).

Neurons in the ventral division of the medial geniculate body show shorter latencies than those in other divisions (Rouiller et al. 1981; Calford 1983; Edeline et al. 1999). However, more recent studies in the mouse and guinea pig thalamus show that neurons in the medial division have shorter latencies than those in the ventral division of the medial geniculate body (Anderson et al. 2006; Anderson and Linden 2011). This evidence strongly supports the shorter latencies in neurons of the anterior auditory field than those in the primary auditory cortex (Schreiner and Urbas 1986, 1988; Linden et al. 2003; Rutkowski et al. 2003; Imaizumi et al. 2004; Bizley et al. 2005) because the anterior auditory field receives input not only from the ventral division but also from the medial division. Latency through thalamocortical transformation can be estimated by in vivo simultaneous single-unit recordings and cross-correlation analysis. A maximum peak in the correlogram is shifted to the expected travel and synaptic delay (e.g., a few milliseconds) (Miller et al. 2001). Thus, it is generally believed that latency in thalamocortical transformation is inherited from the thalamus to cortex. However, a recent study using in vivo whole-cell recordings combined with a pharmacological application (to silence cortical activity by a mixture of muscimol and SCH50911) in the rat primary auditory cortex unfolds a different story: difference in latency between the thalamus and cortex is generated by synaptic integration time by excitation and inhibition through corticocortical interactions (Zhou et al. 2012).

References

Anderson LA, Linden JF (2011) Physiological differences between histologically defined subdivisions in the mouse auditory thalamus. Hear Res 274:48–60

Anderson LA, Malmierca MS, Wallace MN, Palmer AR (2006) Evidence for a direct, short latency projection from the dorsal cochlear nucleus to the auditory thalamus in the guinea pig. Eur J Neurosci 24:491–498

Bao S, Chang EF, Woods J, Merzenich MM (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. Nat Neurosci 7: 974–981

Bartlett EL, Smith PH (1999) Anatomic, intrinsic, and synaptic properties of dorsal and ventral division neurons in rat medial geniculate body. J Neurophysiol 81: 1999–2013

Bartlett EL, Wang X (2007) Neural representations of temporally modulated signals in the auditory thalamus of awake primates. J Neurophysiol 97:1005–1017

Bartlett EL, Wang X (2011) Correlation of neural response properties with auditory thalamus subdivisions in the awake marmoset. J Neurophysiol 105:2647–2667

Bartlett EL, Sadagopan S, Wang X (2011) Fine frequency tuning in monkey auditory cortex and thalamus. J Neurophysiol 106:849–859

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. Nature 403:309–312

Bendor D, Wang X (2007) Differential neural coding of acoustic flutter within primate auditory cortex. Nat Neurosci 10:763–771

Bieser A, Muller-Preuss P (1996) Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. Exp Brain Res 108: 273–284

Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. Nature 451: 197–201

Bizley JK, Nodal FR, Nelken I, King AJ (2005) Functional organization of ferret auditory cortex. Cereb Cortex 15: 1637–1653

Bordi F, LeDoux JE (1994) Response properties of single units in areas of rat auditory thalamus that project to the amygdala. II. Cells receiving convergent auditory and somatosensory inputs and cells antidromically activated by amygdala stimulation. Exp Brain Res 98: 275–286

Bowman EM, Olson CR (1988) Visual and auditory association areas of the cat's posterior ectosylvian gyrus: cortical afferents. J Comp Neurol 272:30–42

Calford MB (1983) The parcellation of the medial geniculate body of the cat defined by the auditory response properties of single units. J Neurosci 3:2350–2364

Chang EF, Bao S, Imaizumi K, Schreiner CE, Merzenich MM (2005) Development of spectral and temporal response selectivity in the auditory cortex. Proc Natl Acad Sci U S A 102:16460–16465

Cheung SW, Bedenbaugh PH, Nagarajan SS, Schreiner CE (2001) Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. J Neurophysiol 85:1732–1749

Clarey JC, Irvine DRF (1990) The anterior ectosylvian sulcal auditory field in the cat: I. An electrophysiological study of its relation to surrounding auditory cortical fields J Comp Neurol 301:289–303

Clascá F, Llamas A, Reinoso-Suárez F (1997) Insular cortex and neighboring fields in the cat: a redefinition based on cortical microarchitecture and connections with the thalamus. J Comp Neurol 384:456–482

Crabtree JW, Collingridge GL, Isaac JT (1998) A new intrathalamic pathway linking modality-related nuclei in the dorsal thalamus. Nat Neurosci 1:389–394

Creutzfeldt O, Hellweg FC, Schreiner C (1980) Thalamocortical transformation of responses to complex auditory stimuli. Exp Brain Res 39:87–104

de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2006) Thalamic connections of the auditory cortex in marmoset monkeys: core and medial belt regions. J Comp Neurol 496:72–96

De Ribaupierre F, Goldstein MH Jr, Yeni-Komshian G (1972) Cortical coding of repetitive acoustic pulses. Brain Res 48:205–225

Dear SP, Simmons JA, Fritz J (1993) A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat. Nature 364: 620–623

Dong C, Qin L, Liu Y, Zhang X, Sato Y (2011) Neural responses in the primary auditory cortex of freely behaving cats while discriminating fast and slow click-trains. PLoS One 6:e25895

Edeline JM, Manunta Y, Nodal FR, Bajo VM (1999) Do auditory responses recorded from awake animals reflect the anatomical parcellation of the auditory thalamus? Hear Res 131:135–152

Eggermont JJ (1998a) Azimuth coding in primary auditory cortex of the cat. II Relative latency and interspike interval representation J Neurophysiol 80:2151–2161

Eggermont JJ (1998b) Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences J Neurophysiol 80: 2743–2764

Eggermont JJ (2000) Sound-induced synchronization of neural activity between and within three auditory cortical areas. J Neurophysiol 83:2708–2722

Eggermont JJ (2001) Between sound and perception: reviewing the search for a neural code. Hear Res 157: 1–42

Ehret G (1997) The auditory cortex. J Comp Physiol A 181:547–557

Furukawa S, Xu L, Middlebrooks JC (2000) Coding of sound-source location by ensembles of cortical neurons. J Neurosci 20:1216–1228

Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J Neurophysiol 32:613–636

Goldstein MHJ, Kiang NYS, Brown RM (1959) Responses of the auditory cortex to repetitive acoustic stimuli. J Acoust Soc Am 31:356–364

Hackett TA (2011) Information flow in the auditory cortical network. Hear Res 27:133–146

Hackett TA, Stepniewska I, Kaas JH (1998) Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 394:475–495

Hackett TA, Barkat TR, O'Brien BM, Hensch TK, Polley DB (2011) Linking topography to tonotopy in the mouse auditory thalamocortical circuit. J Neurosci 31: 2983–2995

Heil P (1997) Auditory cortical onset responses revisited. I First-spike timing J Neurophysiol 77:2616–2641

Huang CL, Winer JA (2000) Auditory thalamocortical projections in the cat: laminar and areal patterns of input. J Comp Neurol 427:302–331

Imaizumi K, Schreiner CE (2007) Spatial interaction between spectral integration and frequency gradient in primary auditory cortex. J Neurophysiol 28: 2933–2942

Imaizumi K, Priebe NJ, Crum PAC, Bedenbaugh PH, Cheung SW, Schreiner CE (2004) Modular functional organization of cat anterior auditory field. J Neurophysiol 92:444–457

Imaizumi K, Priebe NJ, Sharpee TO, Cheung SW, Schreiner CE (2010) Encoding of temporal information by timing, rate, and place in cat auditory cortex. PLoS One 5:e11531

Imaizumi K, Priebe NJ, Cheung SW, Schreiner CE (2011) Spatial organization of repetition rate processing in cat anterior auditory field. Hear Res 280: 70–81

Imig TJ, Morel A (1985) Tonotopic organization in ventral nucleus of medial geniculate body in the cat. J Neurophysiol 53:309–340

Jones EG (2007) The thalamus. Cambridge University Press, Cambridge

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84: 541–577

Kaas JH (2008) The evolution of the complex sensory and motor systems of the human brain. Brain Res Bull 75: 384–390

Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A 97:11793–11799

Kaur S, Lazar R, Metherate R (2004) Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex. J Neurophysiol 91:2551–2567

Kaur S, Rose HJ, Lazar R, Liang K, Metherate R (2005) Spectral integration in primary auditory cortex: laminar processing of afferent input, in vivo and in vitro. Neuroscience 134:1033–1045

Kilgard MP, Merzenich MM (1998) Plasticity of temporal information processing in the primary auditory cortex. Nat Neurosci 1:727–731

Kilgard MP, Merzenich MM (1999) Distributed representation of spectral and temporal information in rat primary auditory cortex. Hear Res 134:16–28

Kishan AU, Lee CC, Winer JA (2008) Branched projections in the auditory thalamocortical and corticocortical systems. Neuroscience 154:283–293

Kishan AU, Lee CC, Winer JA (2011) Patterns of olivocochlear branches. Open J Neurosci 1:1–7

Knight PL (1977) Representation of the cochlea within the anterior auditory field (AAF) of the cat. Brain Res 130: 447–467

Lam YW, Sherman SM (2005) Mapping by laser photostimulation of connections between the thalamic reticular and ventral posterior lateral nuclei in the rat. J Neurophysiol 94:2472–2483

Lam YW, Sherman SM (2007) Different topography of the reticulothalamic inputs to first- and higher-order somatosensory thalamic relays revealed using photostimulation. J Neurophysiol 98:2903–2909

Lam YW, Sherman SM (2010) Functional organization of the somatosensory cortical layer 6 feedback to the thalamus. Cereb Cortex 20:13–24

Lee CC (2013) Thalamic and cortical pathways supporting auditory processing. Brain Lang 126:22–28

Lee CC, Imaizumi K (2013) Functional convergence of thalamic and intrinsic inputs in cortical layers 4 and 6. Neurophysiology 45:396–406

Lee CC, Sherman SM (2008) Synaptic properties of thalamic and intracortical inputs to layer 4 of the first- and higher-order cortical areas in the auditory and somatosensory systems. J Neurophysiol 100:317–326

Lee CC, Sherman SM (2011) On the classification of pathways in the auditory midbrain, thalamus, and cortex. Hear Res 276:79–87

Lee CC, Winer JA (2005) Principles governing auditory cortex connections. Cereb Cortex 15:1804–1814

Lee CC, Winer JA (2008a) Connections of cat auditory cortex: I. Thalamocortical system J Comp Neurol 507: 1879–1900

Lee CC, Winer JA (2008b) Connections of cat auditory cortex: III. Corticocortical system J Comp Neurol 507: 1920–1943

Lee CC, Winer JA (2011a) Convergence of thalamic and cortical pathways in cat auditory cortex. Hear Res 274: 85–94

Lee CC, Winer JA (2011b) A synthesis of auditory cortical connections: thalamocortical, commissural, and corticocortical systems. In: Winer JA, Schreiner CE (eds) The auditory cortex. Springer, New York, pp 147–170

Lee CC, Imaizumi K, Schreiner CE, Winer JA (2004) Concurrent tonotopic processing streams in auditory cortex. Cereb Cortex 14:441–451

Lee CC, Kishan AU, Winer JA (2011) Wiring of divergent networks in the central auditory system. Front Neuroanat 5:46

Liang L, Lu T, Wang X (2002) Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J Neurophysiol 87:2237–2261

Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM (2003) Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. J Neurophysiol 90:2660–2675

Liu BH, Wu GK, Arbuckle R, Tao HW, Zhang LI (2007) Defining cortical frequency tuning with recurrent excitatory circuitry. Nat Neurosci 10:1594–1600

Lomber SG, Malhotra S (2008) Double dissociation of 'what' and 'where' processing in auditory cortex. Nat Neurosci 11:609–616

Lu T, Wang X (2000) Temporal discharge patterns evoked by rapid sequences of wide- and narrowband clicks in the primary auditory cortex of cat. J Neurophysiol 84: 236–246

Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat Neurosci 4:1131–1138

Ma L, Tai X, Su L, Shi L, Wang E, Qin L (2013) The neuronal responses to repetitive acoustic pulses in different fields of the auditory cortex of awake rats. PLoS One 8:e64288

Malone BJ, Scott BH, Semple MN (2007) Dynamic amplitude coding in the auditory cortex of awake rhesus macaques. J Neurophysiol 98:1451–1474

Malone BJ, Scott BH, Semple MN (2010) Temporal codes for amplitude contrast in auditory cortex. J Neurosci 30:767–784

Mardia KV (1972) Statistics of directional data. Academic, London

Miller LM, Escabí MA, Read HL, Schreiner CE (2001) Functional convergence of response properties in the auditory thalamocortical system. Neuron 32: 151–160

Miller LM, Escabí MA, Read HL, Schreiner CE (2002) Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. J Neurophysiol 87: 516–527

Morel A, Imig TJ (1987) Thalamic projections to fields a, AI, P, and VP in the cat auditory cortex. J Comp Neurol 265:119–144

Morel A, Garraghty PE, Kaas JH (1993) Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. J Comp Neurol 335: 437–459

Mountcastle VB (1997) The columnar organization of the neocortex. Brain 120:701–722

Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM (2002) Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. J Neurophysiol 87:1723–1737

Nieuwenhuys R (2013) The myeloarchitectonic studies on the human cerebral cortex of the Vogt-Vogt school, and their significance for the interpretation of functional neuroimaging data. Brain Struct Funct 218:303–352

Niwa M, Johnson JS, O'Connor KN, Sutter ML (2012) Active engagement improves primary auditory

cortical neurons' ability to discriminate temporal modulation. J Neurosci 32:9323–9334
Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK (2008) A voice region in the monkey brain. Nat Neurosci 11:367–374
Philibert B, Beitel RE, Nagarajan SS, Bonham BH, Schreiner CE, Cheung SW (2005) Functional organization and hemispheric comparison of primary auditory cortex in the common marmoset (*Callithrix jacchus*). J Comp Neurol 487:391–406
Pinault D (2004) The thalamic reticular nucleus: structure, function and concept. Brain Res Brain Res Rev 46: 1–31
Polley DB, Read HL, Storace DA, Merzenich MM (2007) Multiparametric auditory receptive field organization across five cortical fields in the albino rat. J Neurophysiol 97:3621–3638
Poremba A, Malloy M, Saunders RC, Carson RE, Herscovitch P, Mishkin M (2004) Species-specific calls evoke asymmetric activity in the monkey's temporal poles. Nature 427:448–451
Preuss A, Muller-Preuss P (1990) Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. Exp Brain Res 79:207–211
Read HL, Winer JA, Schreiner CE (2001) Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. Proc Natl Acad Sci U S A 98:8042–8047
Reale RA, Imig TJ (1980) Tonotopic organization in auditory cortex of the cat. J Comp Neurol 182:265–291
Recanzone GH, Schreiner CE, Sutter ML, Beitel RE, Merzenich MM (1999) Functional organization of spectral receptive fields in the primary auditory cortex of the owl monkey. J Comp Neurol 415:460–481
Rennaker RL, Carey HL, Anderson SE, Sloan AM, Kilgard MP (2007) Anesthesia suppresses nonsynchronous responses to repetitive broadband stimuli. Neuroscience 145:357–369
Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. Philos Trans R Soc Lond Ser B Biol Sci 336:367–373
Rouiller E, de Ribaupierre Y, Toros-Morel A, de Ribaupierre F (1981) Neural coding of repetitive clicks in the medial geniculate body of cat. Hear Res 5: 81–100
Rutkowski RG, Miasnikov AA, Weinberger NM (2003) Characterisation of multiple physiological fields within the anatomical core of rat auditory cortex. Hear Res 181:116–130
Schreiner CE, Cynader MS (1984) Basic functional organization of second auditory cortical field (AII) of the cat. J Neurophysiol 51:1284–1305
Schreiner CE, Raggio MW (1996) Neuronal responses in cat primary auditory cortex to electrical cochlear stimulation. II Repetition rate coding J Neurophysiol 75: 1283–1300
Schreiner CE, Urbas JV (1986) Representation of amplitude modulation in the auditory cortex of the cat. I. the anterior auditory field (AAF). Hear Res 21:227–241
Schreiner CE, Urbas JV (1988) Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. Hear Res 32: 49–64
Schreiner CE, Winer JA (2007) Auditory cortex mapmaking: principles, projections, and plasticity. Neuron 56: 356–365
Schulze H, Langner G (1997) Periodicity coding in the primary auditory cortex of the Mongolian gerbil (Meriones unguiculatus): two different coding strategies for pitch and rhythm? J Comp Physiol A 181: 651–663
Scott BH, Malone BJ, Semple MN (2011) Transformation of temporal processing across auditory cortex of awake macaques. J Neurophysiol 105:712–730
Sherman SM, Guillery RW (2006) Exploring the thalamus and its role in cortical function. MIT Press, London
Smith PH, Bartlett EL, Kowalkowski A (2007) Cortical and collicular inputs to cells in the rat paralaminar thalamic nuclei adjacent to the medial geniculate body. J Neurophysiol 98:681–695
Smith PH, Uhlrich DJ, Manning KA, Banks MI (2012) Thalamocortical projections to rat auditory cortex from the ventral and dorsal divisions of the medial geniculate nucleus. J Comp Neurol 520:34–51
Suga N (1994) Multi-function theory for cortical processing of auditory information: implications of single-unit and lesion data for future research. J Comp Physiol A 175:135–144
Suga N, Manabe T (1982) Neural basis of amplitude-spectrum representation in auditory cortex of the mustached bat. J Neurophysiol 47:225–255
Suga N, Zhang Y, Yan J (1997) Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. J Neurophysiol 77:2098–2114
Sullivan WE 3rd (1982a) Neural representation of target distance in auditory cortex of the echolocating bat Myotis lucifugus. J Neurophysiol 48:1011–1032
Sullivan WE 3rd (1982b) Possible neural mechanisms of target distance coding in auditory system of the echolocating bat Myotis lucifugus. J Neurophysiol 48: 1033–1047
Ter-Mikaelian M, Sanes DH, Semple MN (2007) Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. J Neurosci 27:6091–6102
Wang X, Merzenich MM, Beitel RE, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol 74:2685–2706
Wenstrup JJ (2005) The tectothalamic system. In: Winer JA, Schreiner CE (eds) The inferior colliculus. Springer, New York
Winer JA (1984) The human medial geniculate body. Hear Res 15:225–247
Winer JA, Larue DT (1996) Evolution of GABAergic circuitry in the mammalian medial geniculate body. Proc Natl Acad Sci U S A 93:3083–3087

Winer JA, Lee CC (2007) The distributed auditory cortex. Hear Res 229:3–13
Winer JA, Miller LM, Lee CC, Schreiner CE (2005) Auditory thalamocortical transformation: structure and function. Trends Neurosci 28:255–263
Yin P, Johnson JS, O'Connor KN, Sutter ML (2011) Coding of amplitude modulation in primary auditory cortex. J Neurophysiol 105:582–600
Zhou Y, Mesik L, Sun YJ, Liang F, Xiao Z, Tao HW, Zhang LI (2012) Generation of spike latency tuning by thalamocortical circuits in auditory cortex. J Neurosci 32:9969–9980

Auditory Transducer, Model

► Cochlear Inner Hair Cell, Model

Auditory-Nerve Response, Afferent Signals

Peter Heil
Systems Physiology of Learning, Leibniz Institute for Neurobiology, Magdeburg, Germany

Definition

Sequences of action potentials (spikes) of individual auditory-nerve fibers (ANFs), the primary auditory afferents, in response to sounds impinging upon the ears.

Detailed Description

Anatomical Foundations

Acoustic information relayed from the inner ear to the central nervous system is encoded in the sequences of spikes produced by (type I) ANFs. In mammals, each ANF contacts only one receptor cell (an inner hair cell, IHC) and is excited by transmitter release events from a single ribbon synapse (Ashmore 2010; Matthews and Fuchs 2010; Chapochnikov et al. 2014). Each IHC has 5–30 ribbons, depending upon species and cochlear location (e.g., Meyer et al. 2009; Zhang et al. 2018). The ANFs innervating a given IHC therefore share some, although not all, functional response properties.

Spontaneous Activity

ANFs produce spikes in the absence of external sound (spontaneous activity). The mean spontaneous rate varies between ANFs, in mammals from near zero up to more than 100 spikes per second (e.g., Liberman 1978; Temchin et al. 2008), even between ANFs innervating the same IHC (Wu et al. 2016). The timing of the spikes of a given ANF during spontaneous activity is highly variable. The spike-count statistics and the distribution and serial correlation of inter-spike intervals in an ANF spontaneous spike train can be understood as the result of excitatory transmitter release events produced by the random depletion and random replenishment of a small number of identical but independent presynaptic release sites at each ribbon, in combination with the ANF's refractory properties (Peterson and Heil 2018).

Driven Activity

Sounds impinging on the ipsilateral ear, when of appropriate spectral composition and amplitude, affect the spiking behavior of ANFs, most often increasing the spike rate. A threshold sound level may be defined at which the driven rate of a given ANF exceeds its spontaneous rate by some criterion. The compound action potential (CAP), a gross stimulus-evoked potential which reflects a weighted sum of ANF responses (Bourien et al. 2014), can be recorded in or near the cochlea, such as at the round window.

Frequency Tuning

Each ANF is tuned to sound frequency and is most sensitive, i.e., threshold is lowest, at a particular frequency (the characteristic frequency, CF) which is determined by the position along the cochlear partition of the IHC which it contacts (cochleotopy). Threshold versus frequency curves (tuning curves) are approximately V-shaped but have a plateau region above CF with very high thresholds (Huang and Olson 2011), and curves for high-CF ANFs exhibit low-frequency tails. The sharpness of tuning is often quantified by

the Q-value, defined as the CF divided by the bandwidth of the tuning curve at some level (e.g., 10 dB) above threshold at CF. Q_{10}-values increase with increasing CF (in cats from about 1 to 10 for CFs from 0.2 to 10 kHz; Pickles 2012) but can reach exceptionally high values (>200) in behaviorally relevant frequency ranges in species such as echo-locating bats.

Sound Level Dependence

With increasing sound level, the mean spike rate of a given ANF increases before saturating at several hundred spikes per second at higher sound levels. The range of sound levels over which the spike rate increases (the dynamic range, DR) varies between ANFs. DR is inversely related to the spontaneous rate of an ANF which in turn covaries with the ANF's sensitivity (e.g., Winter et al. 1990). For a given ANF, DR varies with sound frequency and is largest at CF due to compressive growth of mechanical responses in the inner ear with sound level. For frequencies below CF, where the mechanical responses are linear, the increase of the spike rate can be described by a Hill equation with a Hill coefficient of 3 and with the independent parameter being the sum of the sound amplitude and a baseline (Heil et al. 2011). DR and maximum spike rate also adapt to stimulus statistics (Wen et al. 2009).

Adaptation

Adaptation is also manifest as a decrease in spike rate over time in response to sounds of constant amplitude. Within a few milliseconds, the spike rate drops rapidly and then more gradually. The decrease can be modeled as the sum of multiple exponential decays with different time constants and a fractional power law (Zilany et al. 2009; Bruce et al. 2018). Upon cessation of the sound, the spike rate temporarily decreases below the spontaneous rate before recovering over tens to hundreds of milliseconds.

Phase Locking

In response to low-frequency sounds or broadband sounds containing low frequencies, ANFs exhibit phase locking, i.e., spikes are nonrandomly distributed across the period of a low-frequency component (van der Heijden and Joris 2006; Heil and Peterson 2017). Phase locking is often quantified by the measures of vector strength and of the phase angle of the mean vector. For a given ANF, vector strength (a measure of the degree of phase locking) and phase angle of the mean vector vary with frequency and sound level. Across ANFs, maximum vector strength decreases with increasing frequency in a low-pass fashion, with cutoffs of a few kilohertz, depending on species. Phase locking is also seen in the responses of low-CF ANFs to acoustic clicks, elicited by multiple mechanical responses caused by these brief broadband sounds (Guinan 2012). ANFs also phase lock to the modulation envelope of sinusoidal amplitude-modulated tones and noise (e.g., Michelet et al. 2012).

For more detailed reviews of auditory-nerve responses, see Pickles (2012) and Heil and Peterson (2015, 2017). Several of these properties can be altered by sensorineural hearing loss (for reviews, see Young 2012, Henry and Heinz 2013). Loss of ANFs or of their synapses with IHCs can lead to overt and hidden hearing losses (e.g., Kujawa and Liberman 2009). Cochlear implants function by evoking spiking activity in ANFs.

References

Ashmore J (2010) The afferent synapse. In: Fuchs PA, Moore DR (eds) The Oxford handbook of auditory science: the ear. Oxford University Press, Oxford, pp 260–282

Bourien J, Tang Y, Batrel C, Huet A, Lenoir M, Ladrech S, Desmadryl G, Nouvian R, Puel JL, Wang J (2014) Contribution of auditory nerve fibers to compound action potential of the auditory nerve. J Neurophysiol 112:1025–1039

Bruce IC, Erfani Y, Zilany MSA (2018) A phenomenological model of the synapse between the inner hair cell and auditory nerve: implications of limited neurotransmitter release sites. Hear Res 360:40–54

Chapochnikov NM, Takago H, Huang C-H, Pangršič T, Khimich D, Neef J, Auge E, Göttfert F, Hell SW, Wichmann C, Wolf F, Moser T (2014) Uniquantal release through a dynamic fusion pore is a candidate

mechanism of hair cell exocytosis. Neuron 83:1389–1403
Guinan JJ Jr (2012) How are inner hair cells stimulated? Evidence for multiple mechanical drives. Hear Res 292:35–50
Heil P, Peterson AJ (2015) Basic response properties of auditory nerve fibers: a review. Cell Tissue Res 361:129–158
Heil P, Peterson AJ (2017) Spike timing in auditory-nerve fibers during spontaneous activity and phase locking. Synapse 71:5–36
Heil P, Neubauer H, Irvine DRF (2011) An improved model for the rate-level functions of auditory-nerve fibers. J Neurosci 31:15424–15437
Henry KS, Heinz MG (2013) Effects of sensorineural hearing loss on temporal coding of narrowband and broadband signals in the auditory periphery. Hear Res 303:39–47
Huang S, Olson ES (2011) Auditory nerve excitation via a non-traveling wave mode of basilar membrane motion. J Assoc Res Otolaryngol 12:559–575
Kujawa SG, Liberman MC (2009) Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci 29:14077–14085
Liberman MC (1978) Auditory-nerve responses from cats raised in a low noise chamber. J Acoust Soc Am 63:442–455
Matthews G, Fuchs PA (2010) The diverse roles of ribbon synapses in sensory neurotransmission. Nat Rev Neurosci 11:812–822
Meyer AC, Frank T, Khimich D, Hoch G, Riedel D, Chapochnikov NM, Yarin YM, Harke B, Hell SW, Egner A, Moser T (2009) Tuning of synapse number, structure and function in the cochlea. Nat Neurosci 12:444–453
Michelet D, Kovačić P, Joris PX (2012) Ongoing temporal coding of a stochastic stimulus as a function of intensity: time-intensity trading. J Neurosci 32:9517–9527
Peterson AJ, Heil P (2018) A simple model of the inner-hair-cell ribbon synapse accounts for mammalian auditory-nerve-fiber spontaneous spike times. Hear Res 363:1–27
Pickles JO (2012) An introduction to the physiology of hearing, 4th edn. Emerald Group Publishing Limited, Bingley
Temchin AN, Rich NC, Ruggero MA (2008) Threshold tuning curves of chinchilla auditory-nerve fibers. II. Dependence on spontaneous activity and relation to cochlear non-linearity. J Neurophysiol 100:2899–2906
van der Heijden M, Joris PX (2006) Panoramic measurements of the apex of the cochlea. J Neurosci 26:11462–11473
Wen B, Wang GI, Dean I, Delgutte B (2009) Dynamic range adaptation to sound level statistics in the auditory nerve. J Neurosci 29:13797–13808
Winter IM, Robertson D, Yates GK (1990) Diversity of characteristic frequency rate-intensity functions in Guinea pig auditory nerve fibres. Hear Res 45:191–202
Wu JS, Young ED, Glowatzki E (2016) Maturation of spontaneous firing properties after hearing onset in rat auditory nerve fibers: spontaneous rates, refractoriness, and interfiber correlations. J Neurosci 36:10584–10597
Young ED (2012) Neural coding of sound with cochlear damage. In: Henderson D, LePrell CG (eds) Noise-induced hearing loss: scientific advances. Springer, New York, pp 87–135
Zhang L, Engler S, Koepcke L, Steenken F, Köppl C (2018) Concurrent gradients of ribbon volume and AMPA-receptor patch volume in cochlear afferent synapses on gerbil inner hair cells. Hear Res 364:81–89
Zilany MSA, Bruce IC, Nelson PC, Carney LH (2009) A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. J Acoust Soc Am 126:2390–2412

Augmentation

▶ Short-Term Synaptic Plasticity in Central Pattern Generators

Autoassociative Networks

▶ Olfactory Cortical Associative Memory Models

Automated Parameter Search in Small Network Central Pattern Generators

Tomasz G. Smolinski
Department of Computer and Information Sciences, Delaware State University, Dover, DE, USA

Definition

Automated parameter search in small network central pattern generators (CPGs) involves the use of any methods other than manual (i.e., hand-tuning) to generate or tune sets of

parameters that result in physiologically realistic neuronal models of the CPGs. Such methods include "brute-force" explorations of predefined parameter spaces, as well as various heuristics (e.g., multi-objective evolutionary algorithms) used to arrive at a single or more of viable model parameter combinations.

Detailed Description

Central pattern generators (CPGs) are neural networks that produce rhythmically patterned outputs, without relying on any sensory feedback (Hooper 2001). CPGs drive such critical rhythmic activity as breathing, chewing, swimming, walking, heartbeat control, etc. CPGs have been shown to produce rhythmic outputs akin to normal rhythmic activity patterns, even in isolation from other parts of the nervous system, which makes them popular physiological models. Furthermore, due to their relative simplicity, especially in such model organisms as lobsters, crabs, or leeches, CPGs have also become quite widespread in computational modeling studies of cellular and synaptic properties of individual neurons and small neural networks.

While hand-tuning has been traditionally used in the process of creating CPG neuronal models (e.g., Soto-Treviño et al. 2005), in light of recent advances in the computational capabilities of modern computing systems, which now facilitate the use of more complex neuronal models (i.e., in terms of the number of compartments or free parameters), and allow for the exploration of unprecedentedly large parameter search spaces, this approach has become virtually obsolete. Therefore, automated methods for model parameter search have been lately gaining much attention.

There are basically two approaches to the problem of searching for optimal (i.e., physiologically realistic) sets of parameter values for models of small network central pattern generators: (1) "brute-force" explorations of predefined parameter spaces and (2) explorations utilizing heuristic optimization approaches, such as multi-objective evolutionary algorithms (MOEAs).

"Brute-Force" Explorations of Predefined Parameter Spaces

In the case of "brute-force" explorations of predefined parameter spaces, the study usually starts with a hand-tuned model of the CPG, which serves as the "center" for the parameter search space that is created around it. Then, physiologically realistic ranges for the model parameters (e.g., maximal conductances of membrane and synaptic currents) are chosen, along with the granularity for each of the parameters. The granularity determines how many possible values each of the parameters can assume and does not have to be the same for all the parameters, as some of them will exhibit different sensitivities to changes in their values. In some cases, the first step may be omitted and only the parameters, along with their ranges and granularities, are determined.

After such a grid-based parameter search space has been constructed, all of the possible combinations of the parameter values are simulated and tested for their physiological adequacy (possibly under multiple simulation scenarios, such as spontaneous activity, response to current injections, removal of neuromodulation, etc.). Only those models that match the behavior of the biological CPG, which is determined by means of one or more quantitative or qualitative measures of the CPG's characteristics (e.g., spike height, interspiking interval, burst duration, period, preservation of the phase of the rhythm, etc.), are retained for further analyses. However, the rejected models are also sometimes subjected to examination in order to determine what makes "bad" models unacceptable.

Explorations Utilizing Heuristic Optimization Approaches

In the case of explorations utilizing various heuristic approaches, such as multi-objective evolutionary algorithms, the study usually starts with the determination of the model parameters, along with their ranges and granularities, similarly to the "brute-force" approach, but often on a much larger scale. In other words, while the range in

the "brute-force" approach may reflect a 3- to four-fold variation in the parameter values, and the granularity may allow for five to ten possible values, incorporating up to 20-fold variation with hundreds of possible values for each parameter is not unheard of in a heuristic approach. Since this approach is not tasked with simulation and analysis of all of the possible combinations of values in such created parameter search space, it remains computationally feasible.

Another critical step in this approach is the definition of one or more measures of the given CPG's characteristics that will be used to determine physiological adequacy of the models. While such measures can be virtually identical to the ones used in the "brute-force" approach, the difference lies in the fact that they are being used during the process of model generation itself, rather than at the end to filter out the unwanted models. These measures, in essence, become the fitness functions utilized in the process of optimizing model parameter values to drive it toward generating as many as possible "good" models which match the biological system, while limiting the number of "bad" solutions.

After the model parameters and the corresponding ranges and granularities have been determined, and the appropriate fitness functions (possibly multiple, even conflicting) have been defined, the iterative process of optimization of the model parameter values begins and ultimately yields a collection of physiologically realistic CPG models.

Applications

Most of the hitherto applications of the automated parameter searches in small network central pattern generators have been performed in relatively simple invertebrate CPGs. For example, Doloc-Mihu and Calabrese (2011) utilized the "brute-force" approach to construct a large (on the order of terabytes) database of conductance-based models of the half-center oscillator from the leech heartbeat central pattern generator to determine how neuronal parameters influence the network activity. Using the same approach, Günay and Prinz (2010) utilized a large (20,250,000) database of models of the lobster pyloric network to study calcium sensors for network homeostasis. Smolinski, Prinz et al., used both the "brute-force" approach and multi-objective evolutionary algorithms to study the cellular and synaptic properties of the AB/PD (anterior burster/pyloric dilator) pacemaker kernel in the lobster pyloric network (2006, 2009), as well as the conductance correlations involved in the recovery of bursting after neuromodulator deprivation (Shim et al. 2012; Malik et al. 2013).

Cross-References

- ▶ Neuronal Model Databases
- ▶ Neuronal Parameter Space Exploration

References

Doloc-Mihu A, Calabrese R (2011) A database of computational models of a half-center oscillator for analyzing how neuronal parameters influence network activity. J Biol Phys 37:263–283

Günay C, Prinz AA (2010) Model calcium sensors for network homeostasis: sensor and readout parameter analysis from a database of model neuronal networks. J Neurosci 30(5):1686–1169

Hooper SL (2001) Central pattern generators. In: Encyclopedia of life sciences. Wiley, Hoboken

Malik A, Shim K, Prinz AA, Smolinski TG (2013) Multi-objective evolutionary algorithms for analysis of conductance correlations involved in recovery of bursting after neuromodulator deprivation in lobster stomatogastric neuron models. BMC Neurosci 14(1): P370

Shim K, Prinz AA, Smolinski TG (2012) Analyzing conductance correlations involved in recovery of bursting after neuromodulator deprivation in lobster stomatogastric neuron models. BMC Neurosci 13(1): P37

Smolinski TG, Prinz AA (2009) Computational intelligence in modeling of biological neurons: a case study of an invertebrate pacemaker neuron. In: Proceedings of international joint conference on neural networks, Atlanta, pp 2964–2970

Smolinski TG, Soto-Treviño C, Rabbah P, Nadim F, Prinz AA (2006) Analysis of biological neurons via modeling and rule mining. Int J Inf Technol Intell Comput 1(2):293–302

Soto-Treviño C, Rabbah P, Marder E, Nadim F (2005) Computational model of electrically coupled, intrinsically distinct pacemaker neurons. J Neurophysiol 94(2):590–604

Axon Model

► Peripheral Nerve Models

Axon, Modeling

Bruce Graham
University of Stirling, Stirling, UK

Definition

Computational modeling of axons is used to determine the action potential initiation and propagation properties along these long and highly branched structures. Issues under investigation include the site and threshold of action potential initiation, propagation speed in unmyelinated and myelinated axons, and safety factors of propagation through branch points and other geometrical inhomogeneities, such as presynaptic boutons. Modeling allows exploration of the interaction between axonal morphology (diameters and branching structure) and passive and active membrane properties in determining the speed and reliability of action potential propagation. This can include the effects of ion channel noise. Modeling is also used to explore axonal development, including growth cone guidance and branching.

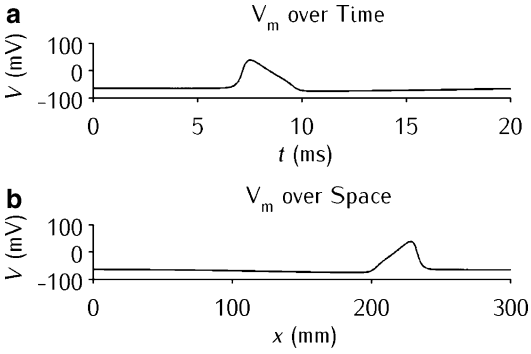
Detailed Description

The axon is the principle communication pathway from one neuron to another. Brief voltage pulses called action potentials (AP) are initiated near the cell body and travel along the axon to reach the many presynaptic boutons. The AP causes neurotransmitter release in a bouton in a stochastic manner, resulting in an electrical response in a target neuron. The reality is more complex than this: AP propagation is not always reliable due to geometrical and electrical constraints in the axon; timing of AP arrival at boutons may vary, having implications for information processing in target nuclei; and AP shape in boutons will influence transmitter release. The experimental data showing these effects is nicely summarized in Debanne (2004). How theoretical results and computational models have been used to aid our understanding of these processes is described in Segev and Schneidman (1999), a paper that is still highly relevant to the field.

Computational modeling is also making a significant contribution to our understanding of how axons develop, from their differentiation from dendrites at initiation to their guidance to target structures and formation of functional connections. This work is documented by van Ooyen in a recent review (van Ooyen 2011) and in an edited volume (van Ooyen 2003).

Hodgkin-Huxley Model of the Action Potential

Computational modeling of the action potential in axons begins with the work of Hodgkin and Huxley in determining a model based on experimental data of the action potential in the squid giant axon (Hodgkin and Huxley 1952). Their model forms the basis of most subsequent models of action potentials and also of models of current flow through membrane-bound ion channels in general in the nervous system. Based on their experimental data, they formulated models of the electrical currents generated by the flow of sodium (Na) and potassium (K) ions across the membrane. Importantly, these currents are voltage sensitive, so that a membrane depolarization results in an increase in the sodium current, which in turn leads to further depolarization and an increase in the sodium current through a positive feedback mechanism. This activation of the sodium conductance is curtailed by a rapid inactivation. The potassium current also increases with depolarization, but more slowly than the sodium current. Eventually, the potassium current grows large enough that it, along with the inactivation of the sodium current, leads to a repolarization of the membrane. This entire process takes place within a few milliseconds, forming a sharp voltage wave of around



Axon, Modeling, Fig. 1 Action potential (AP) in a long, uniform axon. (**a**) AP over time at 25% along the axon. (**b**) AP over space after 20 ms. Simulation performed using the NEURON simulator

100 mV, which is the action potential (Fig. 1a). The basic model describes changes in the transmembrane voltage (V_m) in an isopotential patch of neuronal membrane as a function of membrane capacitance (C_m), sodium (I_{Na}) and potassium (I_{K}) ionic currents, a nonspecific "leak" current (I_L), and experimentally injected electrode current (I_e; to initiate the action potential). The equations of this model are as follows:

Membrane Voltage

$$C_m \frac{dV_m}{dt} = -g_L(V_m - E_L) - g_{\mathrm{Na}} m^3 h(V_m - E_{\mathrm{Na}}) - g_{\mathrm{K}} n^4 (V_m - E_{\mathrm{K}}) + I_e$$

Sodium Current

$$I_{\mathrm{Na}} = g_{\mathrm{Na}} m^3 h(V - E_{\mathrm{Na}}) \frac{dm}{dt} = \alpha_m(1 - m) - \beta_m m$$

$$\frac{dh}{dt} = \alpha_h(1 - h) - \beta_h h$$

$$\alpha_m = 0.1 \frac{V + 40}{1 - \exp(-(V + 40)/10)} \beta_m = 4 \exp(-(V + 65)/18)$$

$$\alpha_h = 0.07 \exp(-(V + 65)/20) \beta_h = \frac{1}{\exp(-(V + 35)/10) + 1}$$

Potassium Current

$$I_K = g_K n^4 (V - E_K) \frac{dn}{dt} = \alpha_n(1 - n) - \beta_n n$$

$$\alpha_n = 0.01 \frac{V + 55}{1 - \exp(-(V + 55)/10} \beta_n = 0.125 \exp(-(V + 65)/80)$$

Action Potential Propagation

This model can be extended to include action potential propagation along the axon using these membrane active properties (sodium and potassium currents) in a compartmental model of an elongated axon (Fig. 1b). The compartmental model arises as a spatial discretization of the partial differential equation describing spatially extensive current flow along an axon:

$$C_m \frac{\partial V_m}{\partial t} = \frac{d}{4R_a} \frac{\partial^2 V}{\partial x^2} - g_L(V_m - E_L) - g_{\mathrm{Na}} m^3 h(V_m - E_{\mathrm{Na}}) - g_{\mathrm{K}} n^4 (V_m - E_{\mathrm{K}})$$

For the simulation shown in Fig. 1, the axon is a cylinder 300 mm long, with a uniform diameter of 476 μm. It is divided into 500 equal-length computational compartments to obtain the numerical solution. The simulation was performed using the NEURON simulator (Carnevale and Hines 2006) with code derived from that used in Chap. 3 of *Principles of Computational Modelling in Neuroscience* (Sterratt et al. 2011).

Speed and Reliability of Propagation

Axons are rarely simply long, uniform cylinders. Many axons become highly branched as they near their target structures. Diameters can change along

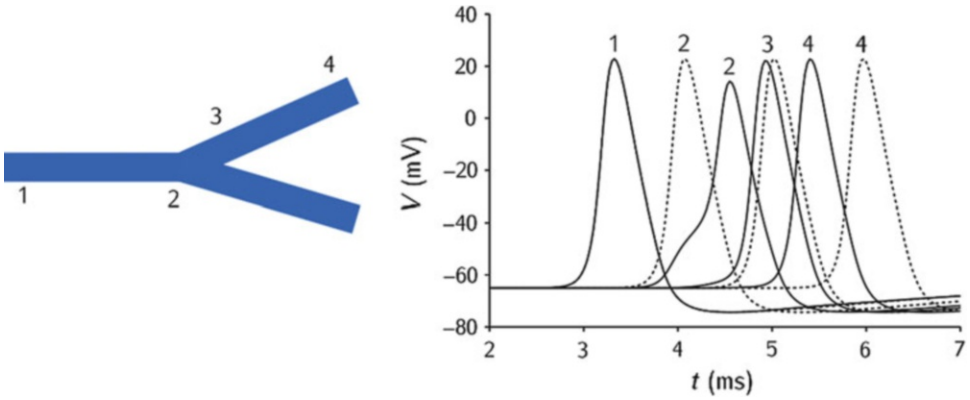
the length of the axon, with a general tapering to thinner diameters toward terminals but with large increases in diameter at presynaptic boutons. Branch points may impose a change in electrical load, depending on the diameters of the daughter branches. The work of Rall and colleagues (see collected papers in Segev et al. 1995) established the important concept of the geometrical ratio (GR) between the diameters of two daughter branches (d_1 and d_2) to that of the parent branch (d_p; section of axon or dendrite closer to the cell body), defined as.

$$\mathrm{GR} = \frac{d_1^{3/2} + d_2^{3/2}}{d_p^{3/2}}$$

Given uniform membrane properties, if GR = 1, the branch point imposes no change in electrical impedance, and the branching can be collapsed to an equivalent uniform cylinder. However, if GR > 1, the branch point imposes an extra electrical load that slows the AP, reduces its height, and can potentially lead to a failure of AP propagation. In contrast, if GR < 1, AP velocity and its height both increase as the branch point is approached, and propagation is reliable. Goldstein and Rall (1974) provided the first comprehensive computational investigation of AP propagation through specified geometrical inhomogeneities. An example is shown in Fig. 2. Here there is a single bifurcation in a long axon. The parent axon (at the left) has a diameter of 1 μm. The two daughter branches have an equal diameter of 2.52 μm, giving GR = 8. The action potential (AP) traces are from the parent axon, branch point, and one daughter branch, as illustrated in the figure. In this configuration (solid lines in time plots), the AP is significantly slowed and reduced in amplitude at the branch point (second trace from left), before quickly recovering and increasing in velocity. If the daughter diameters are reduced to 0.63 μm, then GR = 1, and the configuration is equivalent to a single, uniform axon of 1 μm diameter. In this case, the AP propagates with uniform amplitude and velocity (dashed lines). Note that though the AP at the branch point is significantly delayed when GR = 8, compared to GR = 1, the larger daughter branch diameter results in a higher velocity for the AP once it is through the branch point, and it quickly overtakes the AP traveling in the uniform axon (right-hand solid and dash plots, respectively). Goldstein and Rall (1974) showed that for axons with the same membrane (R_m) and axial (R_a) resistivity, AP speed is constant in units of length constant per unit time, irrespective of diameter (d). The passive length constant (λ) is.

$$\lambda = \sqrt{\frac{dR_m}{4R_a}}$$

This equates to an increase in speed in units of physical length per unit time with an increase in diameter.



Axon, Modeling, Fig. 2 Action potential (AP) in a branched axon. Recording sites illustrated at *left*. Traces of APs over time shown at *right*. *Solid lines*: GR = 8. *Dashed lines*: GR = 1. Simulations performed using the NEURON simulator

Manor et al. (1991) considered AP propagation in realistic axonal branching structures, with a particular emphasis on the delays imposed by geometrical irregularities. This study shows that though delays in AP propagation at individual branch points are at most small (a few tenths of a millisecond), accumulated delays in a complex, branching axon can result in arrival times at boutons differing by a few milliseconds. Delays or speedups from successive, nearby branch points do not sum linearly. In particular, successive summed delays may be supralinear compared to the delays at the individual, isolated branch points.

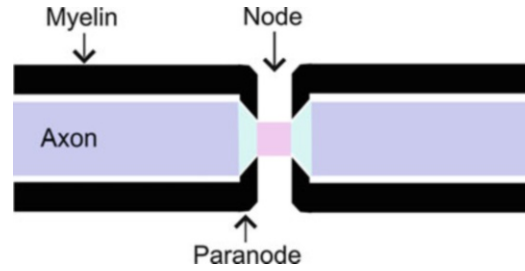
Differential branch point failure, in which the AP fails to propagate along one daughter branch, while successfully propagating along the other, which has been seen experimentally (Debanne 2004), cannot arise due to the branch point geometry alone (Segev and Schneidman 1999). Changes in the active membrane properties are also required and may arise due to differences in ion channel distributions between branches or differential extracellular ion accumulation during successive AP propagations leading to alterations in ion current reversal potentials (particularly E_K; see the Hodgkin-Huxley model current equations) and hence current magnitudes.

In summary, computational modeling has been fundamental to improving our understanding of the characteristics of and constraints to AP propagation along branched and irregular axons.

Myelinated Axon

Propagation speeds along axons containing a continuous distribution of ion channels along their length (so-called unmyelinated axons) are rather modest, on the order of <1 m/s. Most long axons in the nervous system are myelinated, with glial cells providing a high-resistance membrane sheath around the axons for much of their length. Ion channels are largely restricted to the sites of regular breaks in this sheath, known as *nodes of Ranvier*. Action potential propagation speeds are much greater along these axons than along unmyelinated fibers, on the order of 10–100 m/s.

Computational modeling contributes to our understanding of AP propagation along myelinated axons, particularly the implications of the very specific locality of sodium and potassium ion channels at and around the nodes of Ranvier (Fig. 3). Sodium channels are located at the node, with potassium channels largely restricted to the paranodal region.



Axon, Modeling, Fig. 3 Myelinated axon. Structural details around a node of Ranvier

In a simplistic approach, myelination may be modeled as an increased passive resistance and a decreased capacitance of the internodal (myelinated) axonal membrane, along with non-uniform distribution of active sodium and potassium channels. However, the narrow extracellular (periaxonal) space afforded by the myelination, plus the spatial separation of sodium and potassium channels, may require the modeling of ion concentrations and the electro-diffusion of ions intra- and extracellularly, to fully capture the action potential propagation properties of the myelinated axon (Nygren and Halter 1999).

Models show that excessive accumulation of potassium in the periaxonal space following high-frequency repetitive activity can lead to action potential failure (Parnas and Segev 1979). Models are also used to explore the impact of loss of myelin, as occurs in diseases such a multiple sclerosis (Coggan et al. 2010). Loss of myelin leads to a slowing of action potential propagation or even action potential failure altogether.

Ion Channel Noise

Noise in the nervous system comes in many different forms (Faisal et al. 2008), with a major source being the stochastic opening and closing of ion channels in neuronal membrane. The thinness of axons makes them vulnerable to this form

of noise. Computational models in which the deterministic ion channel models of Hodgkin and Huxley are replaced by equivalent stochastic models are used to investigate the impact of ion channel "noise" on action potentials (Goldwyn and Shea-Brown 2011).

The standard approach to turning a deterministic model of the action potential, such as the Hodgkin-Huxley model, into an equivalent stochastic model is to treat the ionic conductance in a patch of membrane of a particular ionic species as being a function of the number of open ion channels for that species. Firstly, the dynamics of conductance change in the deterministic model can be rewritten in the form of a kinetic scheme, which is interpreted as the concentration of ion channels in different states, one of which is an "open" state and provides the conductance. An example of such a scheme for the potassium conductance in the Hodgkin-Huxley model is.

$$[n_0] \overset{4\alpha_n,\beta_n}{\rightleftharpoons} [n_1] \overset{3\alpha_n,2\beta_n}{\rightleftharpoons} [n_2] \overset{2\alpha_n,3\beta_n}{\rightleftharpoons} [n_3] \overset{\alpha_n,4\beta_n}{\rightleftharpoons} [n_4]$$

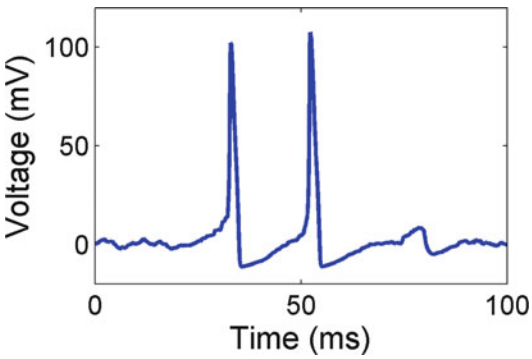
Now the total conductance for potassium in a patch of membrane is.

$$g_K(v,t) = \gamma_K[n_4]$$

where γ_K is the conductance of a single open potassium channel.

This becomes a stochastic model if the patch of membrane is assumed to contain N ion channels and the rates of the kinetic schemes become transition probabilities between states for individual channels, resulting in continuous-time Markov chain. The stochastic model approaches the deterministic model as N goes to infinity. However, for finite numbers of all ion channels, N, the stochastic model can exhibit significantly different behavior from the deterministic model.

Ion channel numbers in a patch of axon may be such that the variation in membrane potential from the opening and closing of individual ion channels may be apparent, particularly near the threshold of action potential initiation where the number of open ion channels is small (Schneidman et al. 1998). This can lead to stochasticity in the generation of action potentials, as illustrated in Fig. 4.



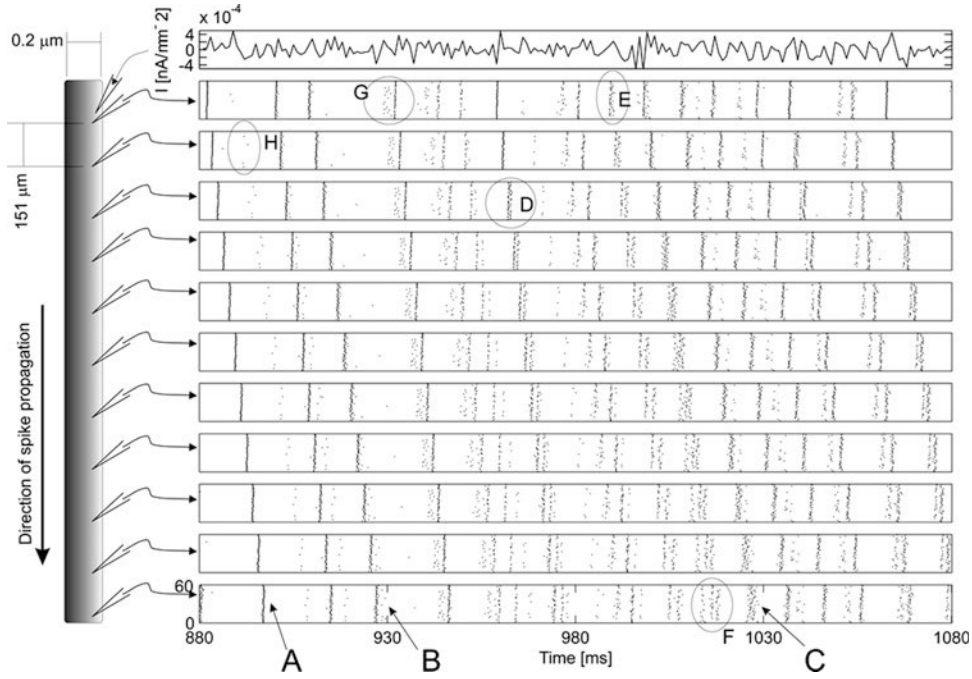
Axon, Modeling, Fig. 4 Spontaneous generation of action potentials in a stochastic Markov model of a 100 μm^2 patch of membrane containing 6000 sodium and 1800 potassium channels. Simulation performed using the MATLAB code of Goldwyn and Shea-Brown (2011) as available in ModelDB (accession number 138950)

Fluctuations in ion channel opening can advance or retard action potential initiation and even lead to failures in propagation, though propagation is more reliable than initiation (Faisal and Laughlin 2007). Multiple trials of a stochastic model of axon stimulation by a fluctuating current are shown in Fig. 5.

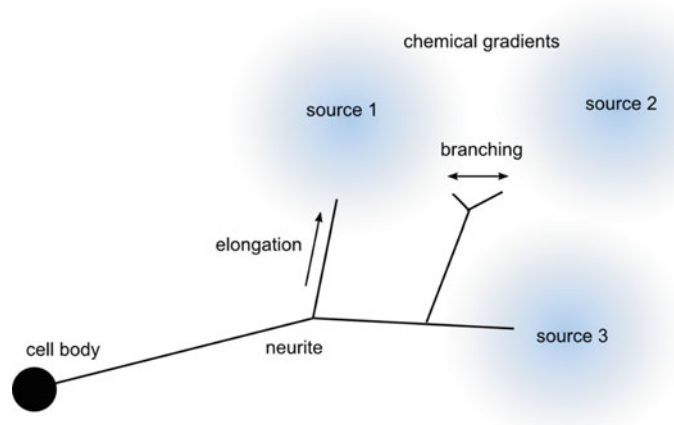
Simulating the Markov model for all ionic species is computationally demanding. Other ways of incorporating membrane noise, which involve some stochastic description of the noise contribution of channel fluctuations, have been tried and can provide a good approximation to the full Markov description (Goldwyn and Shea-Brown 2011).

Axon Development

A completely different class of model is used to explore the development of axons. A major issue here is how an axon reaches its target nucleus and forms synapses therein. Models of axon guidance make explicit the interaction between extracellular attractive and repulsive cues and intracellular growth mechanisms in the growth cone and trailing axon. This requires modeling some form of extracellular environment that contains physical barriers to axon growth and diffusible or membrane-bound molecules acting as attractive and repulsive cues (Fig. 6). Krottje and van



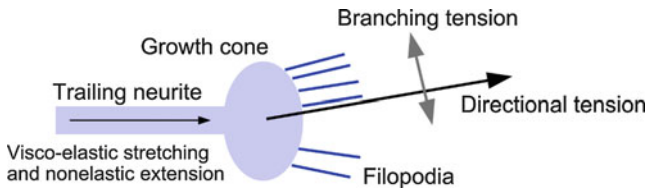
Axon, Modeling, Fig. 5 Action potential initiation and propagation due to fluctuating current stimulation in a stochastic membrane model of a squid-type axon with 0.2 µm diameter. The topmost row shows the stimulus current. Below, each row contains spike raster plots of 60 repeated trials recorded at equally spaced axonal positions. Data is extracted from 10-s trials (Fig. 1 from Faisal and Laughlin (2007), reproduced with permission according to the Creative Commons Attribution License)



Axon, Modeling, Fig. 6 Neurite outgrowth in an external environment containing three sources of a diffusible attractive chemical. (Reproduction of Fig. 10.10 from Sterratt et al. (2011), with permission of the authors)

Ooyen (2007) provide a suitable mathematical description of such an environment in the form of partial differential equations (PDEs) and quasi-steady-state approximations solved across a spatial grid. Axons themselves may not play a space-filling role in the environment.

Models concerned with the details of axon growth and branching, on the other hand, need to include the mechanisms of intracellular signaling, the viscoelastic properties of the axon, and the spatial extent of growth cones (Fig. 7).



Axon, Modeling, Fig. 7 Axonal pathfinding. Detection of chemoattractants in the external environment by filopodia produces tension on the growth cone in particular directions. The growth cone will turn toward and grow along the dominant direction. If similar forces are exerted on opposite sides of the cone, the tension may be enough to split the cone into two, leading to the formation of daughter branches (Reproduction of Fig. 3 from Graham and van Ooyen (2006), with permission according to the Creative Commons Attribution License)

Computational simulation environments, including CX3D (Zubler and Douglas 2009) and NETMORPH (Koene et al. 2009), provide facilities for modeling aspects of axonal and dendritic development and the subsequent formation of neural networks. A review of modeling and computer simulation techniques for neuronal development is provided by Graham and van Ooyen (2006).

Cross-References

► Hodgkin-Huxley Model

References

Carnevale NT, Hines M (2006) The NEURON book. Cambridge University Press, Cambridge

Coggan JS, Prescott SA, Bartol TM, Sejnowski TJ (2010) Imbalance of ionic conductances contributes to diverse symptoms of demyelination. Proc Natl Acad Sci U S A 107:20602–20609

Debanne D (2004) Information processing in the axon. Nat Rev Neurosci 5:304–316

Faisal AA, Laughlin SB (2007) Stochastic simulations on the reliability of action potential propagation in thin axons. PLoS Comput Biol 3(5):e79

Faisal AA, Selen LPJ, Wolpert DM (2008) Noise in the nervous system. Nat Rev Neurosci 9:292–303

Goldstein SS, Rall W (1974) Changes in action potential shape and velocity for changing core conductor geometry. Biophys J 14:731–757

Goldwyn JH, Shea-Brown E (2011) The what and where of adding channel noise to the Hodgkin-Huxley equations. PLoS Comput Biol 7(11):e1002247

Graham BP, van Ooyen A (2006) Mathematical modelling and numerical simulation of the morphological development of neurons. BMC Neurosci 7(Suppl 1):S9

Hodgkin A, Huxley A (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol Lond 117:500–544

Koene RA, Tijms B, van Hees P, Postma F, de Ridder A, Ramakers GJ, van Pelt J, van Ooyen A (2009) NETMORPH: a framework for the stochastic generation of large scale neuronal networks with realistic neuron morphologies. Neuroinformatics 7(3):195–210

Krottje J, van Ooyen A (2007) A mathematical framework for modelling axon guidance. Bull Math Biol 69(1):3–31

Manor Y, Koch C, Segev I (1991) Effect of geometrical irregularities on propagation delay in axonal trees. Biophys J 60:1424–1437

Nygren A, Halter JA (1999) A general approach to modeling conduction and concentration dynamics in excitable cells of concentric cylindrical geometry. J Theor Biol 199:329–358

Parnas I, Segev I (1979) A mathematical model for conduction of action potentials along bifurcating axons. J Physiol Lond 295:323–343

Schneidman E, Freedman B, Segev I (1998) Ion channel stochasticity may be critical in determining the reliability and precision in spike timing. Neural Comput 10:1679–1703

Segev I, Schneidman E (1999) Axons as computing devices: basic insights gained from models. J Physiol Paris 93:263–270

Segev I, Rinzel J, Shepherd GM (eds) (1995) The theoretical foundation of dendritic function: the collected papers of Wilfrid Rall. Bradford Book/MIT Press, Cambridge, MA

Sterratt D, Graham B, Gillies A, Willshaw D (2011) Principles of computational modelling in neuroscience. Cambridge University Press, Cambridge

van Ooyen A (ed) (2003) Modeling neural development. Bradford Book/MIT Press, Cambridge, MA

van Ooyen A (2011) Using theoretical models to analyse neural development. Nat Rev Neurosci 12:311–326

Zubler F, Douglas R (2009) A framework for modeling the growth and development of neurons and networks. Front Comput Neurosci 3:25

Axon-Glial Signaling

► Neuron-Glial Interactions