The auditory novelty system: An attempt to integrate human and animal research

CARLES ESCERAab and MANUEL S. MALMIERCAc,d
aInstitute for Brain, Cognition and Behavior (IR3C), University of Barcelona, Catalonia, Spain
bCognitive Neuroscience Research Group, Department of Psychiatry and Clinical Psychobiology, University of Barcelona, Catalonia, Spain
cAuditory Neurophysiology Laboratory, The Institute of Neuroscience of Castilla y Leon (INCyL), University of Salamanca, Salamanca, Spain
dDepartment of Cell Biology and Pathology, The Medical School, University of Salamanca, Salamanca, Spain

Abstract

In this account, we attempt to integrate two parallel, but thus far, separate lines of research on auditory novelty detection: (1) human studies of EEG recordings of the mismatch negativity (MMN), and (2) animal studies of single-neuron recordings of stimulus-specific adaptation (SSA). The studies demonstrating the existence of novelty neurons showing SSA at different levels along the auditory pathway’s hierarchy, together with the recent results showing human auditory-evoked potential correlates of deviance detection at very short latencies, that is, at 20–40 ms from change onset, support the view that novelty detection is a key principle that governs the functional organization of the auditory system. Furthermore, the generation of the MMN recorded from the human scalp seems to involve a cascade of neuronal processing that occurs at different successive levels of the auditory system’s hierarchy.

Descriptors: Mismatch negativity (MMN), Stimulus-specific adaptation (SSA), Evoked potentials, Single unit recordings, Middle-latency responses (MLR), Auditory brainstem response (ABR), Change detection

The ability to detect the occurrence of unexpected novel stimuli in the acoustic environment is critical for survival, as they might urge a prompt adaptive response. Novelty detection appears as a basic property of the whole nervous system, triggering a cascade of neurological events that engages perception, attention, learning, and memory. For animals in natural environments, such sound noises are made by prey or predators, and communication sounds of conspecifics (e.g., Casseday & Covey, 1996; Fay & Popper, 1994; Malmierca, 2003). For humans, examples of the power of novel events to attract attention are easy to enumerate: let’s say a heartbreaking scream in the middle of the night, the sudden rupture of a glass, or the screeching of a car applying its breaks. One critical function of the auditory brain is hence to alert the organism to sounds that are behaviorally important. Yet, not only largely salient, novel, “unique” stimuli are capable of driving attention involuntarily, but also small stimulus changes occurring occasionally in an otherwise repetitive sound sequence, that is, “deviant,” or contextually novel stimuli, are capable of attracting attention. In order for the nervous system to determine whether a stimulus is novel, there must therefore be some ongoing storage of information about what stimuli have already occurred, as well as an ongoing online comparison of new sounds with previous ones. However, despite the adaptive importance of novelty detection, currently, the underlying neural mechanisms are not fully understood.

The detection of auditory novel events has been associated with a particular brain response extracted from the human electroencephalogram (EEG), the mismatch negativity (MMN; see Figure 1; Näätänen, Gaillard, & Mäntysalo, 1978; for a recent review, see Näätänen, Paavilainen, Rinne, & Alho, 2007) auditory-evoked potential (AEP). MMN is obtained using the auditory oddball paradigm, in which a sequentially repeated stimulus, referred to as the standard, is occasionally replaced by a deviant or novel sound differing in any of its attributes from the repeating one (Näätänen et al., 2007). The MMN is measured as the difference between the AEP elicited in response to the deviant sound and that elicited to the standard sound (or by a respective identical sound that is presented in the role of a standard in a further “reversed” condition; Figure 1A). It usually peaks in humans at 150–200 ms from change onset and has a frontocentral negative scalp distribution, with positive voltages at electrode positions below the Sylvian fissure, indicating generator sources located bilaterally to the supratemporal plane of the auditory cortex (AC; Escera, Alho, Schröger, & Winkler, 2000). A further contribution from prefrontal cortex has
Figure 1. Novelty-related effects derived from the human EEG. A: Typical experimental protocol to obtain the true memory-based comparison deviance effects. The reversed block controls for physical stimulus characteristics when comparing the deviant response to the response obtained to the very same physical stimulus in the role of standard in the reversed block. The control block controls for refractoriness due to stimulus probability when comparing the response to the deviant stimulus versus that elicited to the same stimuli in the control block. B: The auditory evoked potential (AEP) recorded with a broad filter setting (0.5–500 Hz) to the same physical stimulus in the role of deviant (red), standard (blue), and control (black). C: When applying the appropriate filter settings (0.6–35 Hz), the long-latency evoked potential (LAEP) can be elicited, and the difference wave resulting from the subtraction of the standard, or control response from that to the deviant reveals the mismatch negativity (MMN). Notice that the “true” MMN resulting from the control–deviant difference is smaller than that obtained from the subtraction deviant–standard, as this includes an N1-enhancement resulting from refractoriness. D: The enhancement of the Nb by a frequency deviance. The responses were obtained from the same EEG recording when applying a filter setting suited for the middle-latency response (15–200 Hz). Adapted from Grimm et al. (2011).
been suggested by studies of patients with cerebral lesions (Alho, Woods, Algazi, Knight, & Näätänen, 1994) and scalp current density analysis (Deouell, 2007; Giard, Perrin, Pernier, & Bouchet, 1990; Yago, Escera, Alho, & Giard, 2001). Also, the MMN has been recorded in several nonhuman species, with electrodes placed on the scalp, epidurally or subdurally (e.g., Astikainen, Ruusuvirta, Wikgren, & Penttonen, 2006, Astikainen et al., 2011; Javitt, Steinschneider, Schroeder, Vaughan, & Arezzo, 1994; Nakamura et al., 2011; Ruusuvirta, Koivisto, Wikgren, & Astikainen, 2007; Ruusuvirta, Penttonen, & Korhonen, 1998; Tikhomrovov et al., 2008). The interest raised by the MMN to study novelty detection in humans is that (a) it provides a reliable EEG signal of auditory deviance detection (Escera & Grau, 1996; Escera, Yago, Polo, & Grau, 2000), and (b) it can be recorded passively; that is, giving no specific instruction to the subjects, which makes it suitable to study noncooperative populations, such as patients with cerebral lesions or in a coma and even newborns (Näätänen et al., 2011).

A critical aspect in MMN elicitation is that the deviant stimuli occur when no physical standard stimulus is present for comparison, so that the brain’s neurophysiologic response to such rare stimuli requires a comparison memory trace of the preceding repetitive stimulus features, that is to say, a neural representation of the preceding sound regularity (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001; Näätänen & Winkler, 1999; Winkler, Dehnam, & Nelken, 2009). However, this interpretation has been challenged by studies arguing that the MMN might actually reflect an enhancement of the N1 component elicited in response to the physically different features of the deviant stimuli, which occur with a much smaller probability than the standard ones and therefore find feature-specific neural populations on a less refractory state than those responding to the standard, thus raising a much larger N1 (Jääskeläinen et al., 2004). This might be the case, for instance, when the deviant stimulus differs from the standard in a large frequency contrast, as frequency-specific neurons will show attenuated N1 responses to the repeating standard stimulus, while deviants in the oddball sequence will stimulate at least partly a different neural population (due to the tonotopic organization of the auditory system; Malmierca et al., 2008). Similar arguments can be made for other auditory stimulus features. In other words, in the best case there might be an overlap between memory-neuron-based MMN and the refractoriness-based N1 effect. In order to circumvent this conceptual criticism by disentangling the N1-MMN component overlap, the so-called controlled protocol (Figure 1A) has been devised (Schröger & Wolff, 1996). In this controlled protocol, a sequence consisting of a random presentation of different equiprobable sounds allows for comparing the response to a deviant stimulus occurring in an oddball sequence with the response to the corresponding control stimulus, which is physically identical and occurs with exactly the same low probability. Critically, the separation between control stimuli in the relevant stimulus dimension should be at least as large as the separation between the standard and deviant stimuli in the oddball block, so that the responses to the deviant are at least less adapted by other preceding stimuli than in the oddball block. By using the controlled protocol, AEP differences between the deviant stimuli occurring in an oddball sequence and the same stimuli occurring in a nonregular acoustic context cannot reflect the consequences of different states of refractoriness, therefore strongly supporting the violation of stimulus regularity extracted from the acoustic background. Using this controlled protocol, genuine novelty detection based on memory-trace comparison (“true” MMN) has been shown for location (Schröger & Wolff, 1996), pitch (Jacobsen & Schröger, 2001), intensity (Jacobsen, Horenkamp, & Schröger, 2003), and duration (Jacobsen & Schröger, 2003) deviants. Yet, while the controlled paradigm controls for stimulus probability, the stimuli in the oddball block follow a rule of regularity, while those in the control block do not. This has recently been sorted out by devising a new “cascade” sequence in which tone frequency is rising and falling again (Ruhnau, Herrmann, & Schröger, 2012), yielding similar results to those obtained with the controlled block, but controlling additionally for regularity.

Genuine memory-based novelty detection can also be inferred, by means of the MMN, in the cases in which the acoustic regularity is not based on the repetition of a simple acoustic feature, but on some more complex aspects of the auditory input. For example, MMN has been recorded to phonetic stimuli only for individuals in whom the phonetic contrast corresponded to a perceptual phonetic category, that is, belonging to their language repertoire, but not on those for whom the same physical contrast corresponded to a foreign language (Näätänen et al., 1997; Winkler et al., 1999). Moreover, the MMN can be elicited in response to violation of even more complex types of acoustic regularity (see Picton, Alain, Otten, Ritter, & Achim, 2000), such as the one defined by a particular combination of two simple acoustic features (e.g., location and frequency; Gomes, Bernstein, Ritter, Vaughan, & Miller, 1997), by the alternating of two different tones (Alain, Woods, & Ogawa, 1994), or by the frequency relationship between two tones within a pair (Saarinen, Paavilainen, Schönä, Tervaniemi, & Näättänen, 1992). In this latter case, the regularity is established by the contingency rule between successive stimuli, for example, when the frequency of the subsequent tone is predicted by the duration of the previous (Bendixen, Prinz, Horváth, Trujillo-Barreto, & Schröger, 2008; see also Paavilainen, Arajärvi, & Takegata, 2007; Paavilainen, 2013, for a recent review). These more complex kinds of regularity, and particularly those defined by the relationship of stimulus features that evolve in time but vary along that feature dimension, had led some authors to propose that the auditory cortex implements preattentive cognitive operations to make predictions about the near feature, a kind of “primitive intelligence” in audition (Näätänen et al., 2001; Näätänen, Astikainen, Ruusuvirta, & Huotilainen, 2010), which is present already at birth (Carral et al., 2005).

Overall, the MMN component of the AEP seems to reflect a general uniform mechanism of novelty detection that can be distinguished from the effects of pure refractoriness of neurons, one that acts (at latencies above 100–150 ms in humans) for deviances violating regularities of different levels of complexity ranging from very simple ones, like the frequent repetition of a sound, up to much more complex or even abstract relationships in the acoustic environment (Näätänen et al., 2001; Picton et al., 2000). According to recent theoretical views, the MMN is therefore reflecting the violation of a neural representation of the regularity implicit in the past auditory stimulation (Winkler, 2007; Winkler et al., 2009). These representations encode rules extracted from the regular intersound relationships, which are mapped to the concrete sound sequence by finely detailed auditory sensory information. Auditory events are then compared with temporally aligned predictions drawn from the regularity representations (predictive models), and the observable MMN response reflects a process of updating the representation of those detected regularities whose prediction was mismatched by the acoustic input (Winkler, 2007). The model suggests that the mechanisms of the auditory deviance detection system serve to sound organization in the brain, so that the MMN-generating process provides the basis of temporal grouping, a
crucial step towards the formation of auditory objects, as supported by studies showing that subjective sound perception can be predicted from the MMN amplitude (Amenedo & Escera, 2000; Täkinen, May, Reinikainen, & Näätänen, 1994). MMN has been used to study the neural correlates of stream segregation as an index of perceptual segregation of sequentially presented sounds (Sussman, Ritter, & Vaughan, 1999; Winkler, Takegata, & Sussman, 2005). Although recent progress has been recently made towards the identification of a direct neural correlate of these sensory memory trace representations (i.e., the repetition positivity [RP]; Baldeweg, 2007; Costa-Faidella, Baldeweg, Grimm, & Escera, 2011; Costa-Faidella, Grimm, Slab, Díaz-Santaella, & Escera 2011; Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005), a clear picture of how this auditory deviance-detection system in humans is implemented within the auditory nervous system is lacking.

Stimulus-Specific Adaptation as the Mechanism of Novelty Detection

A step forward towards the understanding of the neural mechanisms of novelty detection in the auditory system has been recently provided by studies of single-unit recordings in anesthetized animals. For instance, Ulanovsky, Las, and Nelken (2003) first described the existence of neurons in A1 of the cat that signal the occurrence of deviant sounds while reducing their responses to repeated ones (Figure 2). (In animal research studies, the deviant stimulus is often referred to as the “rare” or the “oddball” stimulus; however, for descriptive purposes and clarity here, we will use the term deviant when referring to both animal and human studies.)

The specific adaptation to repeated sounds, which does not generalize to other sounds, is referred to as stimulus-specific adaptation (SSA). These neurons may be considered as “novelty” neurons. In their study, multunit and single-unit responses were recorded in cat’s primary auditory cortex (A1) and in the auditory division of the thalamus (the medial geniculate body; MGB) to auditory stimuli presented in an oddball arrangement similar to that used in human MMN studies. These authors manipulated both the probability of the standard/deviant stimuli (ratios of 90/10, 70/30, and a control in 50% of cases), the intensity and the normalized frequency difference (Δf) between the two tones (Δf = 0.37, 0.10, 0.04, defined as (Δf = (f2 − f1)/ (f1 × f2)^1/2)). Their main finding was that certain types of neurons in A1, but not in the MGB, exhibit strong SSA; that is, they stopped firing after a few repetitions of the standard tone, but nevertheless showed robust responses to the deviant stimuli. The authors suggested that these novelty neurons would be the single-neuron correlate of MMN, and thus of sensory memory. Moreover, given the cortical location of these units, the authors went further to claim that the origin of MMN would be above the thalamus, in line with the mainstream corticocentric view in cognitive neuroscience (Parvizi, 2009; Waal & Ferrari, 2010).

Since the publication of the original study by Ulanovsky and colleagues in 2003, subsequent studies have extended and completed our current knowledge of SSA in the auditory cortex (Fishman & Steinschneider, 2012; Nelken, Yaron, Polterovich, & Hershenhoren, 2013; Pienkowski & Eggermont, 2009; Szymanski, Garcia-Lazaro, & Schnupp, 2009; Taaseh, Yaron, Polterovich, & Nelken, 2011; von der Behrens, Baurle, Kossel, & Gaese, 2009; Yaron, Hershenhoren, & Nelken, 2012).

Yet, more recent findings corroborate the earliest description of novelty units by Bivikov (Bibikov, 1977; Bibikov & Soroka, 1979) in the frog *torus semicircularis* (an homologous of the inferior colliculus (IC) in frogs and the recording of multunit activity to consonant-vowel contrasts in the MGB of the guinea pig (Kraus, McGee, Littman, Nicol, & King, 1994; King, McGee, Rubel, Nicol, & Kraus, 1995). Indeed, a series of single-unit recording studies have described neurons in IC (Ayala & Malmierca, 2013; Ayala, Perez-Gonzalez, Duque, Nelken, & Malmierca, 2013; Duque, Perez-Gonzalez, Ayala, Palmer, & Malmierca, 2012; Lumari & Zhang, 2010; Malmierca, Cristau, Perez-Gonzalez, & Covey, 2009; Malone & Semple, 2001; Patel, Redhead, Cervi, & Zhang, 2012; Perez-Gonzalez, Covey, & Malmierca, 2005; Perez-Gonzalez, Hernandez, Covey, & Malmierca, 2012; Ponnath, Hoke, & Farris, 2013; Zhao, Liu, Shen, Feng, & Hong, 2011) and MGB (Anderson, Christianson, & Linden, 2009; Antunes & Malmierca, 2011; Antunes, Nelken, Covey, & Malmierca, 2010; Bäuerle, von der Behrens, Kössl, & Gaese, 2011) of the rat that exhibit SSA (Figure 2), and novelty responses similar in many respects to those found by Ulanovsky et al. (2003) in the cat A1, therefore challenging the attribution of novelty detection to higher cortical regions.

The demonstration that SSA is present subcortically in the IC and MGB (Ayala & Malmierca, 2013) makes interesting a comparison between the properties of SSA between the cortical and subcortical neurons to analyze to what extent the SSA observed in cortical neurons is similar or not to that seen in subcortical neurons (Figure 2). Further, it is interesting in the context of understanding the mechanisms that operate along the auditory pathway to create,

![Figure 2](image-url)
shape, and/or modulate SSA as a build-up process. Subcortical novelty neurons are similar to cortical ones in that the magnitude of the novelty response, that is, the difference response between the deviant and standard stimulus, is positively correlated with the amount of frequency contrast between the deviant and the standard and the stimulus repetition rate (up to 4/s). Moreover, it is negatively correlated with the probability of occurrence of the deviant stimulus. In both the cortical and subcortical neurons, there appears to be a continuum in the amount of frequency contrast to which neurons are sensitive. Another significant similarity between the SSA observed in cortical and subcortical neurons is the rapid time course of adaptation to the standard stimulus. Across subcortical neurons, there was a rapid response decay over the first 10 presentations followed by a slower decay that reached a maximum within 25 presentations in a series of 400 repetitions, while no significant adaptation to the oddball stimulus over the course of all presentations is observed (Malmierca et al., 2009; Antunes et al., 2010; Ayala & Malmierca, 2013; Ulanovsky et al., 2003).

By contrast, SSA in subcortical neurons differs from that in the cortical neurons because most IC and MGB neurons; that is, the subcortical neurons respond to a given frequency presented as the deviant stimulus with a shorter latency than to the same frequency presented as the standard stimulus (Antunes et al., 2010; Ayala et al., 2013; Ayala & Malmierca, 2013; Duque et al., 2012; Malmierca et al., 2009). In contrast, A1 neurons respond to the standard and deviant stimuli with similar latencies (Ulanovsky et al., 2003). A second important divergence is that the difference between the response to the deviant and standard is the long latency in cortical neurons (Ulanovsky et al., 2003; cf. their Figure 2), whereas in the subcortical neurons it usually occurs during the onset portion of the response, and therefore is short latency (i.e., mostly at the beginning of the response only; Malmierca et al., 2009; cf. their Figure 4). Another difference is that there is no correlation between SSA and breadth of frequency tuning in the auditory cortex (Ulanovsky et al., 2003), whereas these measures are strongly correlated in subcortical neurons (Duque et al., 2012). Finally, yet another essential difference in the SSA expressed by cortical and subcortical neurons is the anatomical location of these neurons. Strong and widespread SSA in the cortex is found in A1 (although detailed studies of SSA in nonprimary auditory field of the cortex, i.e., nonlemniscal regions, are pending; Fishman & Steinschneider, 2012), while neurons from the subcortical regions only located in the nonlemniscal divisions of the IC (Figure 2A; Ayala et al., 2013; Duque et al., 2012; Lumari & Zhang, 2010; Malmierca et al., 2009; Perez-Gonzalez et al., 2005; Zhao et al., 2011) and MGB show strong and widespread levels of SSA (Antunes et al., 2010). However, it is worth mentioning that weak levels of SSA have also been found in the ventral division of the MGB, that is, the lemniscal MGB in mouse (Anderson et al., 2009) and gerbil (Bäuerle et al., 2011). Moreover, the dorsal division of the mouse MGB, that is, a nonlemniscal part of the MGB, does not show significant SSA (Anderson et al., 2009). It is unclear the reason for these discrepancies in the MGB since SSA in the IC studies from different groups clearly show the unique relationship between SSA and the nonlemniscal IC regions (Ayala et al., 2013; Ayala & Malmierca, 2013; Duque et al., 2012; Lumari & Zhang, 2010; Patel et al., 2012; Zhao et al., 2011). The use of different animal species, stimulation protocols, analytical tools, or a combination of all of them may account for the controversy of different results reported in the MGB studies.

The similarities and differences highlighted above suggest that, possibly, SSA is a feature detection property that emerges in the IC and is transmitted downstream to the cortex because no SSA has been detected in the earlier auditory nuclei like the cochlear nucleus (Ayala et al., 2013; Ayala & Malmierca, 2013). Perhaps AC neurons are more specialized to integrate multiple properties across multiple time scales and are less specialized for feature detection, including the feature of novelty (Ulanovsky et al., 2003). On the other hand, however, these similarities and differences may also reflect genuine differential processing mechanisms at the midbrain/thalamus and cortical levels and raise the question of whether SSA is due to a bottom-up or top-down process; that is, whether the SSA that we have observed in the IC is created de novo at the midbrain level and transmitted from there to the cortex, as mentioned above, or whether it is due to descending projections from the cortex. It was originally proposed that auditory SSA has a cortical origin and is propagated to subcortical nuclei through direct corticofugal projections (Nelken & Ulanovsky, 2007; Ulanovsky et al., 2003). In fact, it is well known that the AC has a dense projection to the nonlemniscal regions of the MGB and IC (Games & Winer, 1988; Malmierca & Ryugo, 2011; Saldaña, Feliciano, & Murgain, 1996; Winer, 2006; Winer, Diehl, & Larue, 1998; Winer, Larue, Diehl, & Hefti, 2001), so it is possible that SSA in those regions could reflect, or be enhanced by, descending corticofugal input. Ulanovsky et al. (2003) argued that the long latency of cortical SSA indicates that intracortical processing must be responsible for SSA. In contrast, for most subcortical neurons, the difference between the response to the standard and deviant was maximal at the beginning of the response (i.e., onset portion; Duque et al., 2012; Malmierca et al., 2009). These findings argue against the idea that SSA in the IC is imposed by the descending projections that the IC receives from the AC. The same applies to the MGB, which receives its ascending inputs from the IC and in addition to a much stronger direct descending projection than the IC.

The most likely scenario is that SSA is generated at the IC through intrinsic circuitry (see above), transmitted to the thalamic MGB for further processing, but modulated by descending cortical input so that the system includes both bottom-up and top-down processing. Fishman and Steinschneider (2012) also suggested that SSA in A1 may be partly inherited from subcortical inputs rather than generated de novo at the cortex since SSA in these subcortical auditory nuclei is observed within the earliest onset responses (Anderson et al., 2009; Antunes et al., 2010; Duque et al., 2012; Malmierca et al., 2009). Accordingly, these synchronized short-latency responses would be expected to produce enhanced initial synaptic activation within thalamocortical input layers of A1. However, prominent differences between deviant and standard responses are primarily observed for later cortical activity rather than for the initial current sink in L3 (Fishman & Steinschneider, 2012), suggesting that intracortical processing significantly amplifies SSA originating at subcortical levels (Szymanski et al., 2009). Moreover, recordings of local field potentials in A1 of the rat in response to standard and oddball tones and subjected to current source density analysis suggest that SSA in IC, MGB, and cortex might be arising independently of each other (Szymanski et al., 2009). Indeed, in a recent study, Malmierca and colleagues (Anderson & Malmierca, 2013; Antunes & Malmierca, 2011) analyzed whether or not SSA in the MGB and IC is dependent on the AC for its generation. The authors reversibly deactivated the AC by the cooling technique (Lomber & Malhotra, 2008; Lomber, Payne, & Horel, 1999) and recorded the changes in SSA sensitivity exhibited by MGB (Antunes & Malmierca, 2011) and IC (Anderson & Malmierca, 2013) neurons. At the population level,
the main finding was that deactivation of the AC did not eliminate SSA exhibited by MGB or IC neurons. A generalized decrease in firing rate to both the deviant and standard stimuli was observed during the cooling condition, but the responses were still higher to the deviant stimuli. Thus, the deviant salience in the IC was preserved even after the deactivation of cortical inputs. At the single neuron level, there were differences between the effects that AC deactivation exerted on subcortical neurons. While most MGB neurons were insensitive to the cortical deactivation, a few neurons in the IC seemed to be affected. Overall, the existence of sets of IC neurons differentially affected by the AC deactivation suggests that the AC may relay SSA to selected neurons, but that the main function of the projections from the AC is to exert a modulatory gain control over the level of adaptation for the majority of neurons in both the IC and MGB neurons (Antunes & Malmierca, 2011, 2013). Thus, SSA persisted in the majority of neurons regardless of the lack of functional corticofugal feedback in both the MGB and IC, suggesting that SSA is inherited through lower input channels in a bottom-up manner and/or generated de novo at each level of the auditory pathway (Antunes & Malmierca, 2011). It should be noted, however, that cortical cooling in both the Antunes and Malmierca (2011) MGB study and the Anderson and Malmierca (2013) IC study was unilateral, unlike cortical cooling in the Lomber studies mentioned above, which was bilateral. It is possible that unilateral cortical cooling could produce disinhibition of the contralateral (uncooled) cortex, and activation of subcortical structures through increased descending cortical input on the contralateral side. Although this is unlikely, it might be a possible explanation for the variability of subcortical response changes observed with unilateral cortical cooling.

Regarding the neurophysiologic mechanisms generating SSA, previous studies (e.g., Kraus et al., 1994; Ulanovsky et al., 2003, 2004; Reches & Gutfriend, 2008) have discussed in detail the possible involvement of intrinsic membrane properties (Gollisch & Herz, 2004; Sanchez-Vives, Nowak, & McCormick, 2000a, 2000b; Abbott, Varela, Sen, & Nelson, 1997), synaptic depression, and facilitation (Abbott et al., 1997; Wehr & Zador, 2005), or synaptic inhibition (Tan, Zhang, Merzenich, & Schreiner, 2004). Recent studies in cortical neuron explants (Eytan, Brenner, & Marom, 2003; Wallach, Eytan, Marom, & Meir, 2008) produced an analog of SSA using an oddball design of electrical stimulation of the cortical neuron network in vitro. Most importantly, Eytan et al. (2003) showed that this selective enhancement of the response to the deviant was abolished by blocking GABAergic inhibition with bicuculline. The IC is a major center for the convergence of both excitatory and inhibitory inputs and for combination of information across frequency-specific channels, especially in the nonlemnical regions (Malmierca, Blackstad, & Olsen, 2011). Inhibitory neurotransmission in the IC is mediated by GABAergic and glycinegic receptors, and GABAergic inputs to the IC come from several sources (Malmierca et al., 2003; Malmierca & Hackett, 2010). In a recent study (Perez-Gonzalez et al., 2012), the GABA A receptor antagonist gabazine was applied microiontophoretically to study the role of GABAergic neurotransmission in the generation and/or modulation of SSA. The firing rate of neurons exhibiting SSA was recorded before, during, and after the drug injection. The results demonstrated that the response magnitude, discharge pattern, and latency remained distinct for the deviant and standard stimuli. The main finding was that the blockade of GABA A receptors modified the temporal dynamics of SSA, and although the adaptation was generally reduced, the blockade of GABA A-mediated inhibition did not abolish SSA completely. Hence, Perez-Gonzalez and colleagues (Perez-Gonzalez et al., 2012; Perez-Gonzalez & Malmierca, 2012) concluded that GABA A-mediated inhibition acts as a gain control mechanism that enhances SSA by controlling the neuron’s gain and responsiveness (a similar situation occurs for the MGB; Duque, Malmierca, & Caspary, 2013). A potential scenario to produce SSA would be through multiple frequency inputs in which some synapses undergo adaptation and others do not (Eytan et al., 2003). These possibilities are obviously not mutually exclusive, and could operate in an interactive fashion. Recent studies describe models based on the SSA results published in the physiology literature in which SSA arises through selective adaptation to the frequently occurring inputs (Mill, Coath, Wennekers, & Denham, 2011a, 2011b, 2012). Garagnaini and Pulvermüller (2011) used a neural model mimicking the cortical anatomy of sensory and motor areas and their connections to explain brain activity indexing auditory change and memory access. These authors showed that neuronal adaptation and local inhibition of cortical activity can explain aspects of change detection as observed when a repeated unfamiliar sound changes in frequency, but the brain dynamics elicited by auditory stimulation with well-known patterns (such as meaningful words) cannot be accounted for on the basis of adaptation and inhibition alone.

A further characteristic of neurons exhibiting SSA is that they necessarily need to integrate information about recent stimulus history in order to respond more strongly to a novel stimulus. Therefore, SSA might represent the short-term memory trace that determines the response of the neuron to subsequent stimulation (Gutfreund, 2012; Jääskeläinen et al., 2004; Nelken et al., 2013; Netser, Zahar, & Gutfriend, 2011; Reches & Gutfriend, 2008). The MGB study by Antunes et al. (2010) has demonstrated that a polynomial scale-invariant model can explain a high proportion of the adaptation to the standard of the MGB neurons. Such a power law model may indicate that adaptation occurs over a range of time scales (Drew & Abbott, 2006). Indeed, SSA in A1 neurons has several time scales concurrently, spanning many orders of magnitude, from hundreds of milliseconds to tens of seconds (Ulanovsky et al., 2004), as also shown for human AEPs (Costa-Faidella, Grimm & Mülling, 2011). SSA was therefore proposed as a candidate neuronal mechanism for auditory sensory memory as reported in MMN studies (Haenschel et al., 2005; Jääskeläinen et al., 2004; Nelken & Ulanovsky, 2007). A recent study based on neuronal recordings and evoked local field potentials (eLFP) in the awake rat found enhanced responses to deviants in eLFP but did not find the latent deviant response component that would have been the equivalent to the human MMN (von de Behrens et al., 2009). But more recent studies suggest that there is equivalent epidural MMN-like potentials to high frequency sounds that encode deviance in an analogous way to the human MMN (Astikainen et al., 2011; Nakamura et al., 2011) and that the SSA in the AC has one of the truly distinguishing characteristics of MMN, that is, true deviance detection (Taaseh et al., 2011). However, this is in contrast to Farley, Quirk, Doherty, and Christian (2010), who found that SSA, unlike MMN, is NMDA-independent, and Fishman and Stein Schneider (2012) whose results obtained in the awake monkey reflect SSA rather than deviance detection per se and led them to conclude that the neural mechanisms of deviance detection likely reside in cortical areas outside of A1. (It should be noted that detailed studies of SSA have been made only on A1 and are pending in other nonprimary cortical fields of AC.) Furthermore, a memory-based mismatch response to speech sounds (ba, da, ga) has also been reported in rats (Ahmed et al., 2011) and, interestingly, also hippocampal responses behave similarly (Ruusuvirta, 2013).
Bridging the Gap Between the MMN and SSA

The studies reviewed so far have provided some hints on how the auditory system is organized to detect novel events in the acoustic environment, but we are still lacking a clear unitary picture of these separated lines of research. For instance, in addition to the obvious neurophysiologic differential nature between the activity recorded in single units (i.e., action potentials) and that of evoked and event-related brain potentials recorded from the scalp (postsynaptic activity over hundreds or thousands of neurons), there are remarkable latency differences between the firing onset of most of the novelty units (at about 20–30 ms from stimulus onset in rats; cf. Perez-Gonzalez et al., 2005; Ulanovsky et al., 2003) and the peak (even onset) latency of the human MMN (circa 150 ms; Näätänen, 1992). Moreover, whereas the MMN has been claimed to be of cortical origin since its early times (e.g., Näätänen, 1992), the discovery of novelty neurons in the IC and MGB responding to simple-feature deviants challenges this claim, although, higher-level computational neurons, located in cortex might be involved for encoding more complex types of regularity.

In order to integrate all the above-reviewed findings into a testable conceptual framework, here we suggest that novelty detection is a pervasive property of the auditory system, expanding from lower levels along the auditory pathway in the brainstem to higher-order areas of the cerebral cortex. Furthermore, we suggest that this functional property of the auditory system is organized in a hierarchical manner, so that more complex levels of regularity will be encoded in higher levels along the novelty system’s hierarchy. To challenge this so-called auditory novelty system, a series of recent studies conducted by Escera and colleagues made use of the whole human AEP, which is a complex response composed of three groups of well-characterized waveforms (Picton, 2010; Picton, Hillyard, Krausz, & Galambos, 1974): the auditory brainstem response (ABR), the middle latency response (MLR), and the long-latency AEP, which includes the MMN. At present, it is well established that the successive waves of the ABR ranging from 1–10 ms from sound onset (waves I, II, III, IV, V, and A) originate from lower-to-upper structures in the subcortical auditory pathway starting from the auditory nerve (wave I-II) up to the IC (waves V-A; Stochard, Stochard, & Shabroough, 1979). On the other hand, the MLR is characterized by a sequence of waveforms in the range 12–50 ms, labeled as N0, N0, N, Pa, and Nb (sometimes Pb, equivalent to P50, is included; Picton, 2010), and represent the earliest cortical responses to a sound. Indeed, a magnetoencephalographic (MEG) study has located the cerebral generators of the P0 waveform, peaking at 16–19 ms, to primary auditory cortex, whereas subsequent waveforms had their sources located in secondary auditory areas (Yvert, Crouzeix, Bertrand, Seither-Preisler, & Pantev, 2001).

According to our hypothesis, novelty responses to simple-feature deviants should be detected at the level of early MLR components originating from A1 (and paralleling the findings of novelty units of the cat’s A1; Ulanovsky et al., 2003), and even lower in the hierarchy, possibly at the level of waves A and V of the ABR, presumably originating from IC (Stochard et al., 1979) and paralleling the findings of novelty units in this midbrain structure of the rat (Duque et al., 2012; Lumani & Zhang, 2009; Malmierca et al., 2009; Patel et al., 2012; Perez-Gonzalez et al., 2005). On the other hand, according to our hierarchical model, however, novelty responses to violations of more complex regularities should be expected only in the higher levels of the AEP, that is, in the latency range of the MMN, as described in previous studies (Alain et al., 1994; Gomes et al., 1997; Näätänen et al., 2001; Saarinen et al., 1992).

The results of this series of studies have started to support our model (see Grimm & Escera, 2012, for a preliminary overview). Frequency deviants occurring among a series of repetitive standard tones elicited an enhancement in the latency range of the Nb component of the MLR, that is, circa 40 ms from deviance onset, when compared to the same physical stimulus presented in the role of standard, and even when the very same stimulus occurred among other equally rare stimuli arranged in the controlled protocol (Figure 1D; Alho, Grimm, Mateo-León, Costa-Faidella, & Escera, 2012; Althen, Grimm, & Escera, 2013; Grimm, Escera, Slabu, & Costa-Faidella, 2011; Leung, Cornella, Grimm, & Escera, 2012). Frequency deviants with less pitch salience, such as broadband noises (see Alho et al., 2012, for discussion), do not elicit Nb enhancements, but a much earlier deviance-related response, by the latency range of the Pa component of the MLR, at circa 30 ms (Slabu, Escera, Grimm, & Costa-Faidella, 2010). However, in this later study, the deviant stimuli failed to elicited any deviance-related response in the latency range of the ABR, that is, by the wave V. Yet a subsequent study that made use of the frequency following response (FFR), which, in contrast to the phasic ABR reflects the tonic brainstem response with bursts of activity matching the repetitive peaks present in the acoustic signal (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010), confirmed the involvement of the human auditory brainstem in novelty detection (Slabu, Grimm, & Escera, 2012). In that study, a deviant consonant-vowel /ba/ syllable presented among a repetitive context of /ba/ syllables elicited a deviance-related effect on the FFR even when compared to a control stimulus (Figure 3; Slabu et al., 2012). Further studies revealed early deviance-related responses as early as 20 ms, by the latency range of the Na component of the MLR for location (Cornella, Leung, Grimm, & Escera, 2012; Grimm, Recasens, Althen, & Escera, 2012) and intensity (Althen, Grimm, & Escera, 2011) deviants. The use of the controlled protocol in the
The auditory novelty system

studies described above is critical, as it supports that the observed effects are not due to mere refractoriness but are reflecting a true memory-comparison process at very short latencies, originating even from the human brainstem.

Another study from our lab made use of MEG to disentangle the contribution of separate regions of the auditory cortex to frequency novelty detection (Recasens, Grimm, Capilla, Nowak, & Escera, in press). The results of this study showed that at early latencies, circa 40 ms from change onset by the Nb component of the MLR, even portions of the primary auditory cortex were involved in deviance detection. At later latencies, circa 100 ms from change onset by the latency range of the N1, more lateral and posterior regions of the right superior temporal gyrus were recruited, in agreement with MEG studies of MMN (e.g., Alho et al., 1998). The results of this study support the involvement of a distributed cerebral network of the auditory cortex, including at least two specific regions operating recurrently at two latency ranges, 40 and 100 ms in novelty detection.

On the other hand, however, the studies that have tried to find early deviance-related correlates of complex-rule violations have yielded negative results. In one such study, the complex rule was defined by the alternating of two different tones, one of 650 Hz and one of 800 Hz presented binaurally, whereas the deviant event was a tone repetition, a paradigm similar to that used by Alain et al. (1994). Critically, a simple-feature deviant was presented by shifting occasionally an alternating tone to one of the ears, by means of an interaural time difference. The results were clear again: both deviant types elicited a sizeable MMN, whereas the simple-feature (location) deviant elicited a deviance-related effect by enhancing the Na component of the MLR, in agreement with Grimm et al. (2012), there were no traces of pattern-violation effects in the MLR range (Cornella et al., 2012). In the same vein, a second study that defined the complex rule by a feature conjunction (cf. Gomes et al., 1997) failed again to provide positive results. In this study, the complex rule was defined by tones sharing the high-frequency/right-location or low-frequency/left-location relationship, whereas deviants were defined by the complementary combinations (Althen et al., 2013). As in the study of Cornella et al. (2012), there was also a condition featuring a simple-feature deviant, that is, frequency, and the results yielded a similar pattern of effects. Again, while the two types of deviances, simple and complex, elicited clear MMNs, only the simple-feature (frequency) deviant elicited an MLR effect, that is, resulted on an enhancement of the Nb component (Althen et al., 2013). From these two studies, we can conclude that early processing stages are involved in encoding simple regularities, while higher levels up in the auditory hierarchy, that is, beyond A1, are required for the encoding of more complex types of regularities. Together with the results reviewed above, they support the view of a hierarchical organization of the auditory novelty system.

Conclusions and Further Directions

The present review was an attempt to integrate within a coherent theoretical framework two parallel but separate lines of research on auditory novelty detection: from one side, human studies of EEG recordings of the MMN; from the other side, animal studies of single-neuron recordings of SSA. The studies reviewed so far, showing the existence of novelty neurons at different levels along the auditory pathway’s hierarchy, together with the recent studies showing human AEP correlates of deviance detection at very short latencies, that is, at around 20–40 ms from change onset, and even originating from the brainstem, are extraordinarily suggestive of two remarkable aspects of the functional organization of the auditory system: First, that novelty detection seems to be a key principle governing the auditory system function; second, that the generation of the MMN recorded from the human scalp might involve a cascade of processes taking place at different levels of the auditory system’s hierarchy, at least when novelty occurs with regard to a single physical attribute of the acoustic input. Moreover, the most recent results investigating deviance-detection to violation of complex regularities, essentially those defined by the relationships between simple features of successive acoustic events, or by combinations of different simple features of the same single event, had led to the proposal that the auditory novelty system should be organized in a hierarchical manner, so that more complex types of regularities, such as those mentioned above, should be encoded, and their corresponding violation detected in the higher levels of the auditory hierarchy. This proposal is actually in line with recent hierarchical models of short-term plasticity in memory and attention networks (Jääskeläinen et al., 2011). The theoretical framework proposed in the present review should foster the establishment of novel synergies between these two so-far-separated fields of endeavor, human and animal research, to the better understanding of the functional organization of the auditory system.
For example, future SSA studies should aim at recording individual neurons beyond A1, first to establish whether it is possible to find SSA in those neurons, and second to disentangle whether or not such neurons would be able to encode complex types of regularities. Of course, such attempts need to be undertaken at subcortical level also, as it is still possible that neurons at lower levels of the hierarchy may still show such complex computational properties. On the other hand, human studies should attempt to relate the early and the late deviance-related responses, that is to say, the relationship of the effects seen by the latency range of the MLR with the MMN. Such an attempt would, of course, help to disentangle the functional role and the dynamics of the corresponding generating systems. A preliminary suggestion in this direction comes from the study by Leung et al. (2012). These authors used the so-called multifeature oddball paradigm (Näätänen, Pakarinen, Rinne, & Takegata, 2004), in which every other stimulus is identical and therefore plays the role of the standard, whereas the interspersed stimuli differ each on a different stimulus feature, for example, frequency, intensity, duration, etc. This paradigm rests on the assumption that individual stimulus features are represented separately in auditory sensory memory and not as gestalts of the combined individual features of the stimulus (Deacon, Nousak, Pilotti, Ritter, & Yang, 1998; Nousak, Deacon, Ritter, & Vaughan, 1996), and therefore the different deviant stimuli act as standards for those features that are not varied on a particular acoustic event. Using this paradigm, it was found that only the frequency deviants elicited an MLR correlate of deviance detection, whereas neither duration, intensity nor location deviants yielded any early deviance-detection effect, while all deviants elicited a significant MMN. This pattern of results suggests that the early systems are unable to represent the individual acoustic features of the stimulus separately, rather as a whole or unitary auditory events, and that it is only later, by the MMN latency range, that the auditory system is able to decompose the auditory stimulus in its individual attributes to keep separate memory traces for its individual features (Leung et al., 2012; Ritter, Deacon, Gomes, Iavitt, & Vaughan, 1995). A mechanism like this would allow the auditory system to track auditory objects that are formed by individual features of consecutive auditory events, and might be on the basis of the encoding of regularities based on feature conjunctions (Gomes et al., 1997).

A further issue that deserves future research attention is that of the relationship between the MMN and the RP, and the corresponding correlates at much shorter latencies. By definition, the MMN is operationally defined as the difference waveform between the deviant and the standard responses. In the same vein, SSA is quantified as common SSA index (CSI), which is derived from the combined spike counting for deviant and standard tones. As the RP is being considered, at least in human research, as the neural correlate of the sensory memory trace of the acoustic regularity, one may expect that the MMN as a measure of the violation of the memory trace should show a similar behavior to RP. However, this has been proved not to be true, as the RP positivity shows short- and long-term repetition effects, whereas the MMN does not show these long-term effects (Costa-Faidella, Grimm et al., 2011), and also the MMN and RP show dissociable effects as a function of the timing information provided by the stimulus sequence (Costa-Faidella, Baldeweg et al., 2011).

Some aspects of the results reviewed here, particularly on the early correlates of human novelty detection, will deserved further attention. For example, two studies attempted to find these correlates by the wave V of the ABR, thought to be generated in IC (Stochard et al., 1979), with negative results (Althen et al., 2011; Slabu et al., 2010). In contrast, another study measuring IC activity through the recording of the FFR did find correlates of true deviance detection in the human auditory brainstem (Slabu et al., 2012). However, the very short latency of wave V, at 5–10 ms, suggests the involvement of the ascending lemniscal portion of the IC, contrasting with the well-established nonlemniscal distribution of the animal novelty neurons in IC (Duque et al., 2012; Malmierca et al., 2009; Perez-Gonzalez et al., 2005). On the other hand, it is interesting to note that these human very early correlates showed what can be considered a “component specificity,” that is, Nb effects for frequency changes, Na effects for location changes, and eventually Na-Pa for intensity changes. Again, the reasons for this dissociation are puzzling. It may be due to the specific parameters of the different experiments involved; alternatively, these effects may point to a component specificity, so that those components showing deviance-related effects may actually reflect the activity of the neural populations encoding for the corresponding specific auditory features. This, of course, will need to be tested in future experiments.

In summary, efforts to integrate human and animal research into a unitary theoretical framework, and preferably into common experimental protocols, such as the one attempted here, should help to deepen our understanding of the functional organization of the auditory system.

References


The auditory novelty system


(Received April 25, 2013; Accepted August 6, 2013)