

Hearing in noisy environments: noise invariance and contrast gain control

Ben D. B. Willmore, James E. Cooke and Andrew J. King

Department of Physiology, Anatomy and Genetics, University of Oxford, Sherrington Building, Parks Road, Oxford OX1 3PT, UK

Abstract Contrast gain control has recently been identified as a fundamental property of the auditory system. Electrophysiological recordings in ferrets have shown that neurons continuously adjust their gain (their sensitivity to change in sound level) in response to the contrast of sounds that are heard. At the level of the auditory cortex, these gain changes partly compensate for changes in sound contrast. This means that sounds which are structurally similar, but have different contrasts, have similar neuronal representations in the auditory cortex. As a result, the cortical representation is relatively invariant to stimulus contrast and robust to the presence of noise in the stimulus. In the inferior colliculus (an important subcortical auditory structure), gain changes are less reliably compensatory, suggesting that contrast- and noise-invariant representations are constructed gradually as one ascends the auditory pathway. In addition to noise invariance, contrast gain control provides a variety of computational advantages over static neuronal representations; it makes efficient use of neuronal dynamic range, may contribute to redundancy-reducing, sparse codes for sound and allows for simpler decoding of population responses. The circuits underlying auditory contrast gain control are still under investigation. As in the visual system, these circuits may be modulated by factors other than stimulus contrast, forming a potential neural substrate for mediating the effects of attention as well as interactions between the senses.

(Received 11 April 2014; accepted after revision 30 May 2014; first published online 6 June 2014)

Corresponding author A. J. King: University of Oxford, Department of Physiology, Anatomy and Genetics, Sherrington Building, Parks Road, Oxford OX1 3PT, UK. Email: andrew.king@dpag.ox.ac.uk

Abbreviations A1, primary auditory cortex; IC, inferior colliculus; MGB, medial geniculate body; PV, parvalbumin; SC, superior colliculus; SOM, somatostatin; V1, primary visual cortex; VIP, vasoactive intestinal polypeptide.

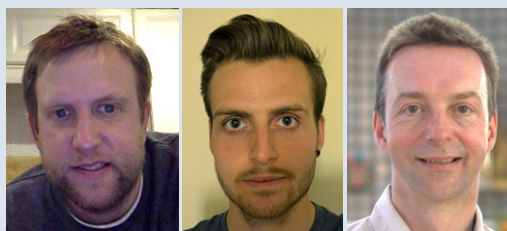
Introduction

Understanding speech is perhaps the most important challenge facing the human auditory system. Listeners with healthy auditory systems can recognize speech with

minimal effort in an extremely wide range of conditions, from a whisper in a silent room to a shout in a raging storm. This belies the difficulty of speech recognition as a computational problem. Despite a huge economic imperative and a proportionately large amount of effort

Ben Willmore is a neurophysiologist and modeller, interested in how neurons in the sensory systems process and represent complex stimuli. He received a BA in Natural Sciences and a PhD in Physiology from the University of Cambridge; his PhD thesis investigated population coding of natural scenes in real and modelled visual cortical neurons. After postdoc work on the visual system at the University of California, Berkeley, he moved to the Auditory Neuroscience Group at the University of Oxford where he records from and builds models of neurons in the auditory system. **James Cooke** is a Wellcome 4 year Neuroscience DPhil student, interested in cortical function and circuit organisation. He completed a BA in Experimental Psychology and an MSc in Neuroscience at Oxford

University and is now based in the Auditory Neuroscience Group for his DPhil. His research involves using electrophysiological and optogenetic techniques to investigate the biophysical basis of cortical computations. **Andrew King** is a Wellcome Principal Research Fellow and Professor of Neurophysiology at the University of Oxford. He received a BSc in Physiology from King's College London and carried out his PhD research at the National Institute for Medical Research. He has held several research fellowships in Oxford and heads the Auditory Neuroscience Group, which combines behavioural, electrophysiological, imaging and computational approaches to investigate the neural basis of auditory perception and how this is shaped by experience.



and investment, computer speech recognition systems are far from perfect, lagging well behind the abilities of an average human listener. This suggests that our understanding of the speech recognition problem has some way to go.

Speech recognition is most difficult when a voice is heard in a noisy environment. The most famous and challenging example of this is the ‘cocktail party problem’ (Cherry, 1953); in a room containing many simultaneous speakers, understanding speech requires us to separate one particular voice from the babble of many superimposed voices, each of which may be perceptually and statistically similar to the voice of interest. But speech recognition is challenging even when the background sounds are not speech. Computer speech recognition systems are far more sensitive than humans to the presence of background noise. Furthermore, people with a wide range of hearing disorders (including age-related hearing loss) find that understanding speech in the presence of background noise is difficult. This is often the case even if hearing aids are used to compensate for their raised thresholds, suggesting that the difficulty they experience in understanding speech in noise may arise from a deficit in central auditory processing (Eggermont, 2014). These findings, therefore, highlight the importance of the background noise problem and suggest that current therapies do not adequately solve it.

Invariance to background noise

It is useful to think of the background noise problem in statistical terms. An auditory scene can be thought of as the sum (superposition) of a signal, S (the sound of interest; for example, speech that we are trying to recognize), and background noise, N . The challenge facing the auditory system is statistically to separate S from N . Of course, the statistics of S and N may vary from moment to moment, so the auditory system must be able to adapt dynamically to changes in these statistics. The cocktail party provides a worst-case scenario, where the statistics of S and N may be very similar to one another. Singling out S in this case is likely to involve highly sophisticated processes involving inference at multiple levels (from sound waveforms to phonemes to grammar), and a complete understanding of this problem is likely to be decades away. However, many commonly occurring situations involve forms of background noise that are more tractable.

Consider the case where S is a human voice (whose frequency content and sound level vary over time) and N is a droning sound (whose frequency content and sound level are constant), perhaps from a fan that is periodically switched on and off. In this case, a relatively simple neural process may be enough to allow the brain to represent S in a way that is invariant to background noise. A first attempt

at solving this problem computationally might involve simply filtering out the sound frequencies of the background noise. This could be done in the brain by silencing neurons that respond to these frequencies. However, many possible background noises, such as a fan or running water, are broadband, meaning that they contain a very wide range of frequencies, overlapping those in a human voice. In such cases, filtering out the noise frequencies would also filter out the voice.

Adaptation to stimulus statistics

The brain has a more effective solution to this problem. Instead of blindly filtering out the frequencies of the background noise, neurons continually adapt their responses to match the statistics of the sounds that are heard (Dean *et al.* 2005, 2008; Baccus, 2006; Nagel & Doupe, 2006; Watkins & Barbour, 2008; Robinson & McAlpine, 2009; Zilany *et al.* 2009; Rabinowitz *et al.* 2011; Wen *et al.* 2012). Consider the effect of background noise on the overall level of a sound. Sound level is generally measured logarithmically (on a decibel scale), reflecting the large dynamic range of the auditory system and the fact that the perceived loudness of a sound generally grows as a logarithm of its magnitude. As the level of the background noise increases, the mean overall sound level (μ) of $S + N$ increases and the variance (σ^2) decreases (Fig. 1; see Rabinowitz *et al.* 2011). If these statistics change sufficiently slowly, it should be possible for neurons to adapt to the statistical changes, reducing their responses to the background sound.

Adaptation to mean sound level has been observed at multiple levels of the auditory system, notably the auditory nerve (Wen *et al.* 2009) and inferior colliculus (IC; Dean *et al.* 2005). When the mean sound level is high, neurons shift their dynamic ranges upwards, so that they are more sensitive to louder sounds. These threshold changes are compensatory, i.e. the changes in neuronal sensitivity tend to compensate for the changes in mean stimulus level, so that the neuronal responses are relatively invariant to changes in background level.

In the primary auditory cortex of the ferret (A1), neurons also show compensatory adaptation to sound level variance, or contrast (Rabinowitz *et al.* 2011). This process is known as contrast gain control. When the contrast of the input to a given neuron is high, the gain of the neuron is low, so that the neuron is relatively insensitive to changes in sound level. When the contrast of the input is low, the gain of the neuron is high, increasing its sensitivity. Thus, the gain of the neuron changes in such a way that it tends to compensate for changes in sound contrast.

The combined effect of adaptation to the mean and contrast of sounds is to minimize the responses of cortical neurons to a statistically stationary background sound, N . The remaining neuronal responses mainly depend on

the signal, S , and are relatively invariant to the contrast of S . This confers a degree of noise invariance directly on the responses of cortical neurons and enables them to represent complex sounds, such as speech, in a fashion that is robust to the presence of background noise (Rabinowitz *et al.* 2013; Mesgarani *et al.* 2014; Fig. 2).

Other processes underlying noise invariance

It is important to note that adaptation to stimulus statistics is not the only strategy used by the auditory system to separate signal from noise. The processes described above work well only for signal and noise combinations with particular characteristics; that is, where the statistics of the noise, N , are constant (or change only slowly over time), but the signal, S , is constantly varying. Under other circumstances (such as the cocktail party problem itself, where multiple voices are present simultaneously), this approach will not be sufficient. The brain must therefore employ additional strategies to separate signal from noise.

Some of these strategies are already understood. It has been shown, for example, that neurons in the midbrain

of songbirds (Woolley *et al.* 2006) and gerbils (Lesica & Grothe, 2008) adapt their modulation tuning preferences in order to reduce their responses to background noise. Neurons in avian auditory cortex acquire noise invariance through tuning for long sounds with sharp spectral structure (Moore *et al.* 2013). Moreover, in humans, a non-linear representation of sound modulations contributes to noise robustness (Pasley *et al.* 2012). Finally, our spatial hearing plays a key role in separating different sound sources that are present simultaneously, subsequently improving our ability to identify them (Yost, 1997; Kidd *et al.* 2005).

It has been shown using magnetoencephalography recordings that the human brain forms separate representations of the voices of simultaneous speakers (Ding & Simon, 2012). It is likely that this separation reflects the action of several of the above processes, plus others which operate at higher levels (for example, taking account of phonemic and grammatical structure), and is used to separate attended speech from the wide variety of background sounds that are encountered in the real world.

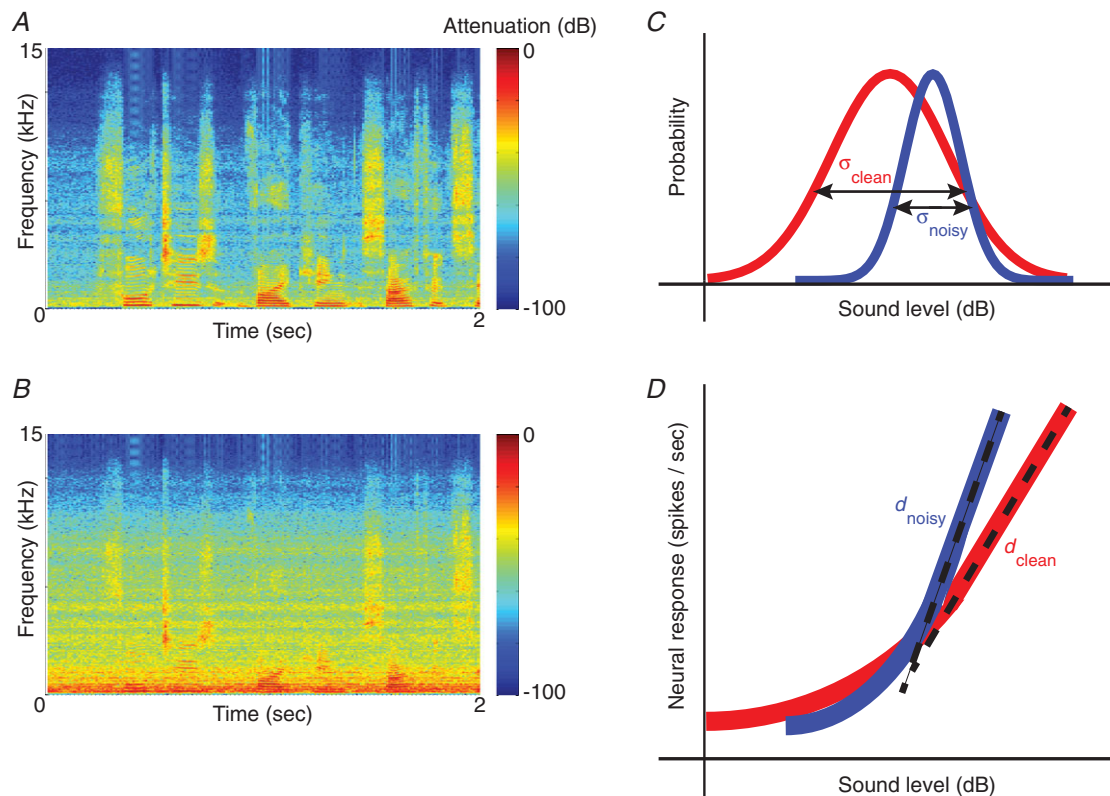


Figure 1. Adaptation to stimulus contrast

A, spectrogram of a sample of human speech (a 'clean' sound), showing how the frequency content of the speech changes over time. *B*, spectrogram of the same speech sample, with noise added (a 'noisy' sound). *C*, idealized sound-level distributions of the clean (red) and noisy (blue) sound, showing that the noise reduces the standard deviation, σ , of the signal. *D*, idealized responses of a neuron with contrast gain control. The neuron adjusts its gain, d , according to the standard deviation of the signal, so that, for both clean and noisy sounds, the dynamic range of the neuron matches the range of stimulus levels encountered.

Contrast gain control in other sensory modalities

The auditory cortex is not the only part of the brain that uses contrast gain control; similar processes have been observed in the retinae of cats (Shapley & Victor, 1978, 1981), salamanders and rabbits (Baccus & Meister, 2002), the primary visual cortex (V1) of cats (Heeger, 1992*a,b*) and primates (Carandini & Heeger, 1994) and the *Drosophila* olfactory system (Olsen *et al.* 2010). Contrast gain control in the visual and auditory cortices seems to behave in similar ways; in both cases, the gain changes can be described by identical equations (Heeger, 1992*a,b*; Rabinowitz *et al.* 2011). In the visual system, contrast is computed locally, so that the gain of each neuron is determined by the contrast of the visual image over an area of the retina close to the receptive field of the neuron (Webb *et al.* 2003), rather than over the entire retinal image. Likewise, in the auditory cortex, the gain of each neuron is determined by the contrast only at sound frequencies that are within the receptive field of the neuron (Rabinowitz *et al.* 2012), suggesting that contrast is calculated in a comparable fashion in these two sensory modalities. Indeed, the similarities between these processes are sufficiently striking that it has been suggested that contrast gain control may be a canonical neural computation (Carandini & Heeger, 2012).

Gain control in multisensory processing

Combining information from different sensory systems can have a profound effect on perception and behaviour, by improving stimulus detection and discrimination,

reducing perceptual uncertainty, and by speeding up reaction times (reviewed by Alais *et al.* 2010). At the neuronal level, the principles underlying multisensory integration have been revealed most clearly in the mammalian superior colliculus (SC), which receives converging visual, auditory and somatosensory inputs and is involved in orienting the eyes and head toward salient sensory cues. The largest gain changes tend to be seen when the individual stimuli are weakly effective in driving the neurons (Meredith & Stein, 1986) and when those stimuli are presented in close temporal and spatial proximity (King & Palmer, 1985; Meredith *et al.* 1987; Meredith & Stein, 1996).

Although these principles also apply in a broad sense to the effects of multisensory stimulation on both the responses of neurons in other parts of the brain and behaviour, it is clear that the manner in which stimuli are combined depends on contextual factors, such as past experience and behavioural relevance, too (van Atteveldt *et al.* 2014). Nevertheless, several attempts have been made to define the computations underlying multisensory integration. On the basis of differences in the way pairs of stimuli within and across sensory modalities interact to determine the responses of SC neurons, Alvarado *et al.* (2007) argued that different rules operate for the integration of unisensory and multisensory cues. More recently, however, the divisive normalization model developed to explain contrast gain control in the visual cortex (Heeger, 1992*a*) has been shown to account for key aspects of multisensory integration, including its dependence on the relative effectiveness and spatial locations of the individual stimuli (Ohshiro *et al.* 2011).

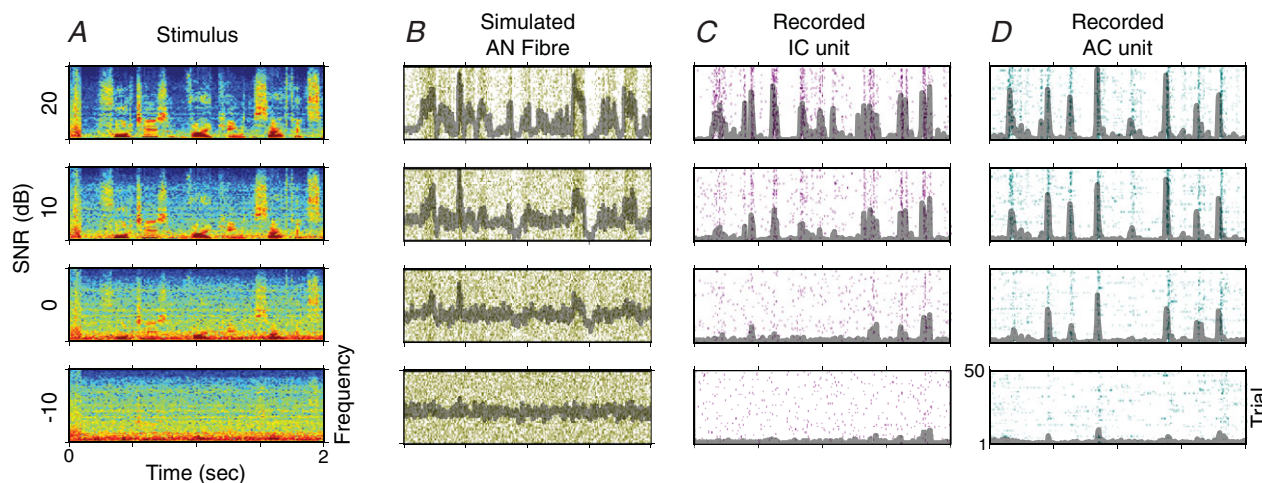


Figure 2. Noise-invariant neural responses

A, spectrograms of a clean sound (a sample of speech, top), embedded in increasing levels of noise (other rows). B, simulated responses of an auditory nerve (AN) fibre to these sounds. The fibre shows clear responses to the clean sound (top), but these are rapidly degraded as more noise is added (other rows). C and D, recorded responses of neurons in ferret inferior colliculus (IC) and auditory cortex (AC) to the same sounds. In IC (C) and especially in AC (D), the lower rows more closely resemble the top row, indicating that neuronal responses in the central auditory pathway are more robust to the addition of noise (i.e. they are more noise invariant). Based on Rabinowitz *et al.* (2013).

Contrast gain control in subcortical and cortical processing

In the visual literature, contrast gain control is frequently referred to as contrast normalization. We have so far avoided using this term to refer to contrast gain control in the auditory system. In general, normalization refers to a complete compensation. For example, Z-scoring is a form of normalization where values are divided by the standard deviation (much like contrast normalization). This division means that Z-scoring completely compensates for the effect of changing the standard deviation, so that Z-scores from different situations can be compared directly.

In the visual system, it has been shown that contrast gain control operates at multiple levels, from retina (Shapley & Victor, 1978, 1981; Baccus & Meister, 2002) to cortex (Heeger, 1992*a,b*). The combined effect of these multiple stages of gain control is to approximate contrast normalization. In the auditory system, however, contrast gain control has not yet been shown to be sufficiently complete to be described accurately as normalization (Fig. 3). In the IC, stimulus contrast affects neuronal response gain, but the gain changes in individual neurons do not compensate for changes in contrast as reliably as those in cortex; different neurons have different strengths and even directions of gain control (Dean *et al.* 2005; Rabinowitz *et al.* 2013). The cumulative effect of these small, variable gain changes in individual IC neurons is to produce a population representation that does show contrast gain control (Dean *et al.* 2005), but at the level of individual neurons the overall effect is not one of uniform contrast normalization.

In ferret A1, most neurons are subject to compensatory gain control, but even here the effect is not strong enough to compensate completely for changes in stimulus contrast. Instead, gain changes compensate for approximately two-thirds of stimulus contrast changes, so that (for most neurons) responses to high-contrast sounds are still somewhat stronger than responses to low-contrast sounds (Rabinowitz *et al.* 2011).

This suggests that consistent, compensatory contrast gain control is not a general property of auditory processing, but is constructed gradually as one ascends the auditory system. Complete contrast normalization may be a desirable property of neural circuits (for example, it would produce more complete noise invariance), so it is possible that higher levels of the auditory cortex may perform gain control that is more fully compensatory, which might accurately be called contrast normalization.

Mechanism of cortical contrast gain control

Little is known about the underlying basis of cortical contrast gain control in the auditory system, so it is pre-

sently unclear whether the same neuronal mechanism applies in all sensory modalities. Thus, it remains to be seen whether contrast gain control at this level is implemented through a canonical cortical processing mechanism or whether different brain regions have independently evolved different mechanisms to implement similar neuronal processing strategies.

A wide variety of possible physiological mechanisms exists, of which two have received particular attention in the literature: shunting inhibition and synaptic depression. Shunting inhibition has long been considered as a mechanism for changing neuronal gain (Carandini & Heeger, 1994; Nelson, 2008). It occurs when neuronal membrane conductance is increased without any change in membrane potential (Borg-Graham *et al.* 1998) and can be the result of either balanced excitation and inhibition or the opening of ion channels with reversal potentials close to the resting potential of the neuron (Reichardt *et al.* 1983).

Shunting inhibition is an appealing mechanism because shunting conductances have a divisive effect on the membrane potential due to Ohm's law (Fatt & Katz, 1953; Coombs *et al.* 1955). Nevertheless, this does not simply translate into a divisive effect on firing rates. In fact, modelling studies have suggested that the effect of shunting conductances on firing rates is subtractive rather than divisive (Gabbiani *et al.* 1994; Holt & Koch, 1997; Capaday, 2002), and this has been confirmed in *in vitro* cortical slice preparations (Ulrich, 2003; Mitchell & Silver, 2003). The consequences of shunting are more complex *in vivo*, however, where locally generated synaptic noise (Borg-Graham *et al.* 1998; Destexhe & Paré, 1999; Destexhe *et al.* 2003) can alter neuronal gain by allowing small inputs to drive the cell (Chance *et al.* 2002), while having little effect on large inputs. This results in shunting conductances altering neuronal gain by scaling the synaptic noise. Highly variable signals are particularly suited to being modulated in this way because this effect depends on the variability of the synaptic input (Mitchell & Silver, 2003), making high-contrast sensory inputs the ideal stimulus for neuronal gain control through shunting inhibition.

A particular class of cortical interneuron appears to be specialized to provide shunting inhibition. These cells have been described anatomically as basket and chandelier cells due to their extensive axonal arbors, which heavily innervate the soma and axon initial segment of pyramidal cells (Markram *et al.* 2004). The inhibition provided by these cells occurs primarily via GABA_A receptors (Klausberger *et al.* 2002), which have a reversal potential close to the resting potential of pyramidal cells. These neurons are therefore ideally placed to provide large, perisomatic shunting conductances required for the effect of this form of inhibition to be divisive. Innervation of the soma also enables these cells to modulate the spiking

responses of pyramidal neurons without affecting their tuning (Isaacson & Scanziani, 2011; Fino *et al.* 2013).

Many basket and chandelier cells have been described physiologically as fast-spiking interneurons (as a result of their tendency to fire bursts of narrow action potentials; Markram *et al.* 2004) and express the calcium-binding protein parvalbumin (Xu *et al.* 2010). In mouse V1, optogenetic manipulation of parvalbumin (PV)-expressing interneurons has been found to alter the gain of visually evoked pyramidal cell spiking responses without affecting the tuning of these cells (Atallah *et al.* 2012; Wilson *et al.* 2012), supporting their possible involvement in contrast gain control.

Synaptic depression, i.e. the reduction in efficacy of a synapse that is repeatedly engaged, can alter neuro-

nal gain in cerebellar granule cells *in vivo* and in simulations of cortical neurons (Abbott *et al.* 1997; Rothman *et al.* 2009) and has also been proposed to account for contrast gain control as well as other physiological properties of cortical neurons (David *et al.* 2009). If synaptic depression of thalamic or intracortical glutamatergic afferents does indeed underlie contrast gain control, the inputs that cause gain changes should be the same as those that excite the neuron. In the auditory cortex, the frequency tuning of gain control is similar, but not identical to the excitatory component of each neuron's receptive field (Rabinowitz *et al.* 2012). This implies that synaptic depression of excitatory inputs alone may not be a sufficient mechanism for gain control.

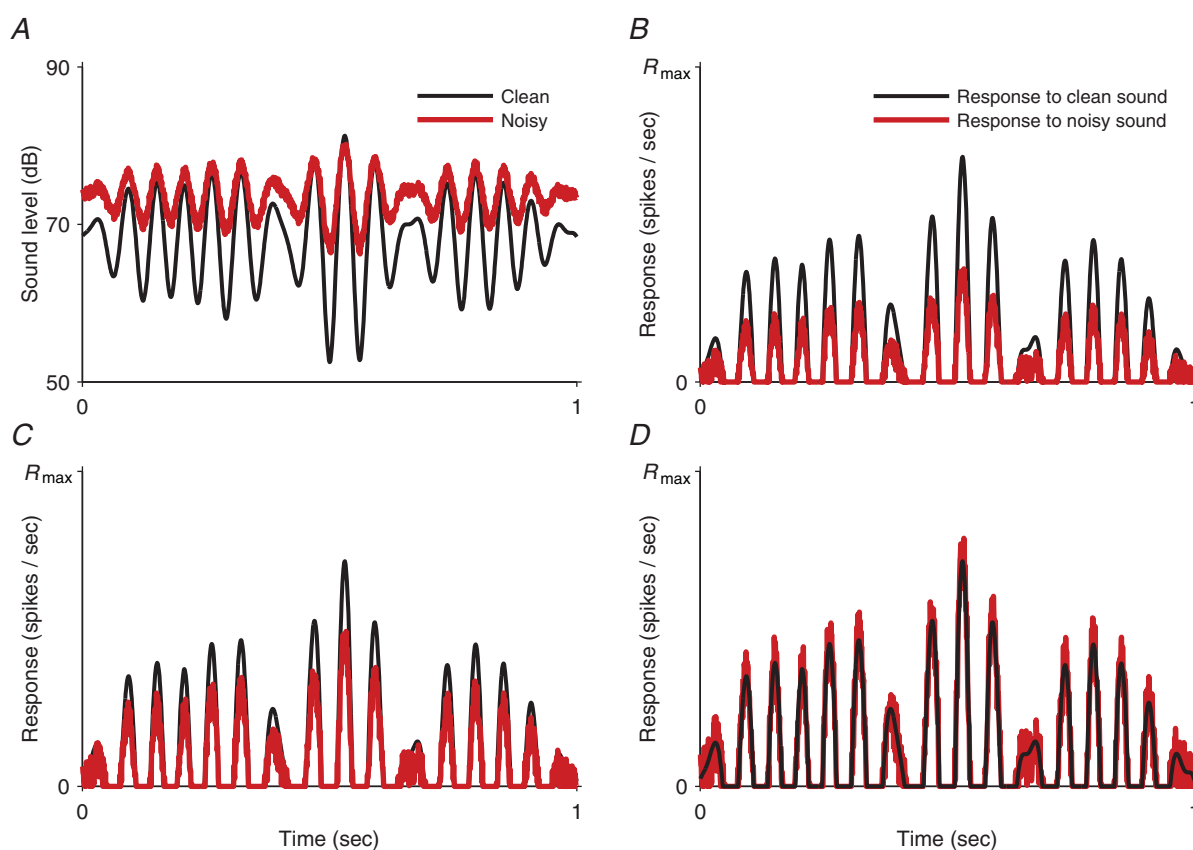


Figure 3. Gain control and normalization

A, the waveform of a clean sound (black line) and the same sound after noise has been added (red line). The effect of the noise is both to increase the mean sound level and to reduce the sound contrast (variance of sound level). B–D, idealized neuronal responses to the clean sound (black) and noisy sound (red). B shows a neuron that adapts its dynamic range to compensate for changes in mean sound level. Due to this adaptation, the neuron does not produce an ongoing response to either the clean or the noisy sound; instead, it responds to deviations of the sound from the mean level. As the neuron does not adapt to compensate for stimulus contrast, the relative strength of the responses depends on whether the sound is clean or noisy. C shows a neuron which also adapts to compensate partly for changes in stimulus contrast. The gain of the neuron is higher for the noisy sound than for the clean sound, and so the strengths of responses to the clean and noisy sounds are more similar than in B. Contrast gain control of this kind is frequently seen in ferret primary auditory cortex. D shows a neuron which completely compensates for changes in both mean sound level and sound contrast. The responses are now very similar, differing only in the fine structure introduced by the noise. Complete compensation of this kind (contrast normalization) has not yet been observed in the auditory system. R_{max} , maximal firing rate.

It remains to be seen whether auditory contrast gain control occurs at a network level or at the level of specific inputs to cortical neurons. The majority of single units in ferret A1 have been shown to exhibit contrast gain control (Rabinowitz *et al.* 2011), suggesting that it may be present in pyramidal cell populations throughout all cortical layers. This may be due to contrast gain control occurring within layer 4, allowing these gain changes subsequently to be inherited by neurons in other cortical layers (Fig. 4). Alternatively, gain may be modulated incrementally within each layer of the cortex. A third possibility is that gain modulation occurs across all cortical layers simultaneously. In mouse V1, corticothalamic layer 6 excitatory neurons have been found to modulate the gain of pyramidal cells in all other cortical layers (Olsen *et al.* 2012) through the recruitment of a translaminal-projecting, fast-spiking interneuron subtype, whose cell bodies also reside within layer 6 (Bortone *et al.* 2014). This deepest cortical layer receives direct thalamic input, which may drive gain changes across the cortical column simultaneously via this intracortical circuit.

Gain control during behavioural tasks may be implemented by modulation of these circuits. In mouse A1, activation of vasoactive intestinal polypeptide (VIP)-expressing interneurons by reinforcement feedback signals can increase the gain of pyramidal cells via inhibition of both somatostatin (SOM) and PV-expressing interneurons (Pi *et al.* 2013). Long-range cortical inputs to mouse barrel cortex have been shown preferentially to target VIP interneurons, which are located in the superficial layers of the cortex, over those expressing PV or SOM, providing a potential route for the gain of cortical pyramidal cells to be modulated by inputs from other areas, including those potentially mediating multisensory or sensorimotor interactions (Lee *et al.* 2013). In mouse A1, layer 1 interneurons can be activated by cholinergic inputs to the cortex (Letzkus *et al.* 2011), resulting in inhibition of layer 2/3 PV-expressing interneurons. VIP-expressing interneurons have been shown to be activated by cholinergic inputs to mouse V1, resulting in an increase in the gain of pyramidal cell responses (Fu *et al.* 2014). This may, therefore, represent a route by which the gain of sensory neurons is modulated by attention or other behavioural states.

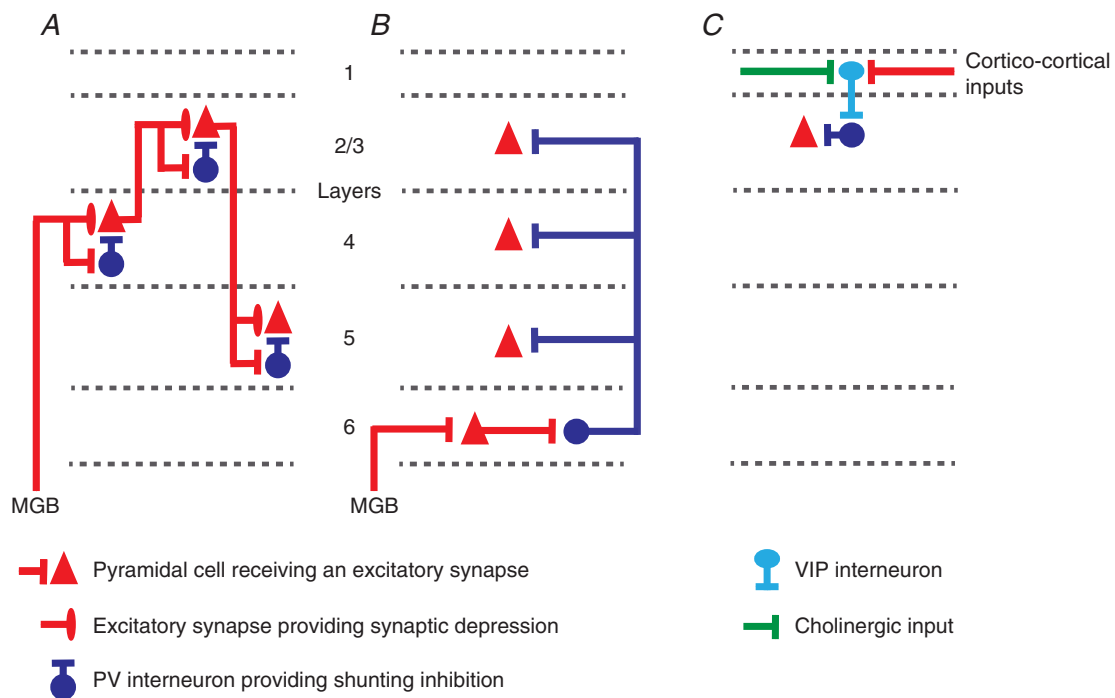


Figure 4. Possible cortical circuits underlying contrast gain control

A, contrast gain control may be implemented within layer 4, by either shunting inhibition or synaptic depression, and these changes may be inherited subsequently by pyramidal cells in supragranular and infragranular layers. Further gain control may also occur within these non-granular layers. B, alternatively, a subpopulation of thalamorecipient layer 6 cells may drive gain changes in all other cortical layers via feedforward inhibition (adapted from Bortone *et al.* 2014). C, vasoactive intestinal polypeptide (VIP)-expressing interneuron activity may modulate pyramidal cell gain by controlling the activity of parvalbumin (PV)-expressing interneurons. The VIP interneurons may be recruited by cholinergic inputs during attention or by corticocortical inputs during multisensory or sensorimotor integration. Abbreviation: MGB, medial geniculate body.

Computational advantages of contrast gain control

We have discussed one main computational advantage of contrast gain control, i.e. that it helps to generate noise-tolerant representations of signals of interest (Rabinowitz *et al.* 2013). There are several more reasons why this is an advantageous coding strategy.

Dynamic range. The simplest of these reasons is that contrast gain control makes efficient use of neuronal dynamic range. The dynamic range is the range of stimulus values encoded by a neuron by a change in its firing rate, which is limited by the maximal rate at which a neuron can fire action potentials (R_{\max} , typically 100–500 Hz).

Imagine a simple, static coding strategy, in which neuronal firing rate, R , is proportional to sound level, L . If neurons employed such a coding strategy, they might encode the loudest sound the animal ever encounters, L_{\max} , with this maximal firing rate, R_{\max} . Under normal circumstances, however, the animal will typically encounter sounds at a fraction of this maximal level. As a result, these static neurons will usually use only a fraction of their dynamic range. Moreover, at any given moment, the range of sound levels that an animal is likely to hear will be correlated; sounds that are nearby in time tend to come from the same sources and are therefore likely to have similar sound levels. This further reduces the proportion of the dynamic range that is used at any given time. As a consequence, small (but potentially important) differences in sound level will be encoded by tiny differences in neuronal firing rate, making them difficult to discriminate.

Contrast gain control provides a better strategy, whereby neuronal gain is continuously adjusted so that the dynamic range of the neurons covers the range of sound levels that the animal is currently experiencing. Thus, the full dynamic range is used to represent the changes in sound level that really occur, improving the accuracy (and the information rate; Laughlin, 1981; Dean *et al.* 2005) of the neuronal representation. Behaviourally, this strategy should also improve an animal's ability to discriminate small changes in sound level when the auditory environment is relatively static.

Redundancy reduction. Another potential advantage of contrast gain control is that it may reduce the redundancy of the neural code. Redundancy is thought to be undesirable in neural representations (Attneave, 1954; Barlow, 1961, 2001). A form of redundancy which is inherent in natural stimuli is that, at any given moment, different stimulus features tend to occur at similar contrasts. For a naïve neural code, this would mean that the responses of sensory neurons would be correlated (Schwartz & Simoncelli, 2001). In principle, contrast gain control ought to reduce these correlations,

however, resulting in a less redundant, sparse code (Field, 1994; Olshausen & Field, 1996) for sensory information. Sparse codes have been observed in the auditory system (Schneider & Woolley, 2013), but it remains to be seen whether contrast gain control is responsible for this.

Population coding. It has been suggested that contrast gain control offers a number of advantages for population coding. It may, for example, have a role in the development of distributed representations whereby neurons become tuned to different features of sensory input (Willmore *et al.* 2012). Contrast gain control also results in population codes that can be decoded easily using a linear classifier (Olsen *et al.* 2010) and that can exhibit winner-take-all behaviour depending on the relative contrast of different stimuli that are presented simultaneously (Busse *et al.* 2009). The latter is a feature that is frequently required by object-recognition algorithms and computation of stimulus saliency (Itti & Koch, 2000, 2001).

Attentional modulation of neuronal gain

The ability to detect a signal against a background of noise depends not only on the way neurons adapt to the stimulus statistics, but also on the level of attention to the task. This has been illustrated by recordings from neurons in the auditory cortex of ferrets trained to discriminate target tones against background noise or to discriminate between different tones or tone complexes (Fritz *et al.* 2007; Atiani *et al.* 2009; Yin *et al.* 2014). These studies have shown that the gain and shape of the spectrotemporal receptive fields of cortical neurons can change within a few minutes of beginning the task in ways that appear to enhance the contrast between the two stimulus categories and presumably, therefore, improve perceptual discrimination. Indeed, the magnitude of the spectrotemporal receptive field plasticity has been found to correlate with task performance, implying a direct relationship between levels of attention and the extent of these physiological changes (Atiani *et al.* 2009).

There are also other examples of selective attention modulating the representation of stimulus information in early sensory cortex in a manner consistent with gain control. Focusing on one speaker in a 'cocktail party' situation leads to enhanced tracking of the attended speech stream in neural activity recorded from the human auditory cortex (Zion Golumbic *et al.* 2013). Likewise, in the visual system, it has been proposed that covert spatial attention (the ability to process information preferentially at a particular point on the retina without moving the eyes) may be subserved by a gain control process, which increases the responses of neurons with receptive fields in the attended location, thereby conferring a processing advantage, and decreases

responses to unattended locations. It is conceivable that attentional gain control (Kerlin *et al.* 2010) might use the same neuronal circuits as contrast gain control; such an arrangement has been proposed in the visual system (Lee *et al.* 1999; Reynolds & Heeger, 2009). In this case, an additional (probably) top-down input to the gain control circuit would be required, modulating the effect of contrast.

Conclusion

Contrast gain control is emerging as an important feature of central auditory processing, as it is in the visual system. It becomes more prevalent as one ascends the auditory pathway, suggesting that one function of higher auditory cortex might be to produce a fully contrast-normalized representation of behaviourally relevant sounds. Such a representation would have numerous computational advantages for the representation of sound, including increased noise invariance, sparse representation and simpler decoding, but additional recording studies will be required to find whether this representation exists in higher auditory areas. More work will also be required to uncover the cellular circuits and synaptic mechanisms responsible for auditory contrast gain control and to determine whether the networks that implement gain control are also those engaged by attentional modulation and during experience-dependent plasticity.

References

- Abbott LF, Varela JA, Sen K & Nelson SB (1997). Synaptic depression and cortical gain control. *Science* **275**, 220–224.
- Alais D, Newell FN & Mamassian P (2010). Multisensory processing in review: from physiology to behaviour. *Seeing Perceiving* **23**, 3–38.
- Alvarado JC, Vaughan JW, Stanford TR & Stein BE (2007). Multisensory versus unisensory integration: contrasting modes in the superior colliculus. *J Neurophysiol* **97**, 3193–3205.
- Atallah BV, Bruns W, Carandini M & Scanziani M (2012). Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73**, 159–170.
- Atiani S, Elhilali M, David SV, Fritz JB & Shamma SA (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* **61**, 467–480.
- Attneave F (1954). Some informational aspects of visual perception. *Psychol Rev* **61**, 183–193.
- Baccus SA (2006). From a whisper to a roar: adaptation to the mean and variance of naturalistic sounds. *Neuron* **51**, 682–684.
- Baccus SA & Meister M (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron* **36**, 909–919.
- Barlow HB (1961). The coding of sensory messages. In *Current Problems in Animal Behaviour*, ed. Thorpe WH & Zangwill OL. Cambridge University Press, Cambridge, UK, pp. 331–360.
- Barlow HB (2001). Redundancy reduction revisited. *Network* **12**, 241–253.
- Borg-Graham LJ, Monier C & Frégnac Y (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature* **393**, 369–373.
- Bortone DS, Olsen SR & Scanziani M (2014). Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex. *Neuron* **82**, 474–485.
- Busse L, Wade AR & Carandini M (2009). Representation of concurrent stimuli by population activity in visual cortex. *Neuron* **64**, 931–942.
- Capaday C (2002). A re-examination of the possibility of controlling the firing rate gain of neurons by balancing excitatory and inhibitory conductances. *Exp Brain Res* **143**, 67–77.
- Carandini M & Heeger DJ (1994). Summation and division by neurons in primate visual cortex. *Science* **264**, 1333–1336.
- Carandini M & Heeger DJ (2012). Normalization as a canonical neural computation. *Nat Rev Neurosci* **13**, 51–62.
- Chance FS, Abbott LF & Reyes AD (2002). Gain modulation from background synaptic input. *Neuron* **35**, 773–782.
- Cherry EC (1953). Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* **25**, 975–979.
- Coombs JS, Eccles JC & Fatt P (1955). The electrical properties of the motoneurone membrane. *J Physiol* **130**, 291–325.
- David SV, Mesgarani N, Fritz JB & Shamma SA (2009). Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by natural stimuli. *J Neurosci* **29**, 3374–3386.
- Dean I, Harper NS & McAlpine D (2005). Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* **8**, 1684–1689.
- Dean I, Robinson BL, Harper NS & McAlpine D (2008). Rapid neural adaptation to sound level statistics. *J Neurosci* **28**, 6430–6438.
- Destexhe A & Paré D (1999). Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *J Neurophysiol* **81**, 1531–1547.
- Destexhe A, Rudolph M & Paré D (2003). The high-conductance state of neocortical neurons *in vivo*. *Nat Rev Neurosci* **4**, 739–751.
- Ding N & Simon JZ (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol* **107**, 78–89.
- Eggermont JJ (2014). *Noise and the Brain*. Academic Press, London.
- Fatt P & Katz B (1953). The effect of inhibitory nerve impulses on a crustacean muscle fibre. *J Physiol* **121**, 374–389.
- Field DJ (1994). What is the goal of sensory coding? *Neural Comput* **6**, 559–601.
- Fino E, Packer AM & Yuste R (2013). The logic of inhibitory connectivity in the neocortex. *Neuroscientist* **19**, 228–237.

- Fritz JB, Elhilali M & Shamma SA (2007). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J Neurophysiol* **98**, 2337–2346.
- Fu Y, Tucciarone JM, Espinosa JS, Sheng N, Darcy DP, Nicoll RA, Huang ZJ & Stryker MP (2014). A cortical circuit for gain control by behavioral state. *Cell* **156**, 1139–1152.
- Gabbiani F, Midtgaard J & Knöpfel T (1994). Synaptic integration in a model of cerebellar granule cells. *J Neurophysiol* **72**, 999–1009.
- Heeger DJ (1992a). Normalization of cell responses in cat striate cortex. *Vis Neurosci* **9**, 181–197.
- Heeger DJ (1992b). Half-squaring in responses of cat striate cells. *Vis Neurosci* **9**, 427–443.
- Holt GR & Koch C (1997). Shunting inhibition does not have a divisive effect on firing rates. *Neural Comput* **9**, 1001–1013.
- Isaacson JS & Scanziani M (2011). How inhibition shapes cortical activity. *Neuron* **72**, 231–243.
- Itti L & Koch C (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res* **40**, 1489–1506.
- Itti L & Koch C (2001). Computational modelling of visual attention. *Nat Rev Neurosci* **2**, 194–203.
- Kerlin JR, Shahin AJ & Miller LM (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party”. *J Neurosci* **30**, 620–628.
- Kidd G, Arbogast TL, Mason CR & Gallun FJ (2005). The advantage of knowing where to listen. *J Acoust Soc Am* **118**, 3804–3815.
- King AJ & Palmer AR (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res* **60**, 492–500.
- Klausberger T, Roberts JDB & Somogyi P (2002). Cell type- and input-specific differences in the number and subtypes of synaptic GABA_A receptors in the hippocampus. *J Neurosci* **22**, 2513–2521.
- Laughlin SB (1981). A simple coding procedure enhances a neuron’s information capacity. *Z Naturforsch C* **36**, 910–912.
- Lee DK, Itti L, Koch C & Braun J (1999). Attention activates winner-take-all competition among visual filters. *Nat Neurosci* **2**, 375–381.
- Lee S, Kruglikov I, Huang ZJ, Fishell G & Rudy B (2013). A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nat Neurosci* **16**, 1662–1670.
- Lesica NA & Grothe B (2008). Efficient temporal processing of naturalistic sounds. *PLoS ONE* **3**, e1655.
- Letzkus JJ, Wolff SBE, Meyer EMM, Tovote P, Courtin J, Herry C & Lüthi A (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* **480**, 331–335.
- Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G & Wu C (2004). Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* **5**, 793–807.
- Meredith MA, Nemitz JW & Stein BE (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci* **7**, 3215–3229.
- Meredith MA & Stein BE (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol* **56**, 640–662.
- Meredith MA & Stein BE (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *J Neurophysiol* **75**, 1843–1857.
- Mesgarani N, David SV, Fritz JB & Shamma SA (2014). Mechanisms of noise robust representation of speech in primary auditory cortex. *Proc Natl Acad Sci U S A* **111**, 6792–6797.
- Mitchell SJ & Silver RA (2003). Shunting inhibition modulates neuronal gain during synaptic excitation. *Neuron* **38**, 433–445.
- Moore RC, Lee T & Theunissen FE (2013). Noise-invariant neurons in the avian auditory cortex: hearing the song in noise. *PLoS Comput Biol* **9**, e1002942.
- Nagel KI & Doupe AJ (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron* **51**, 845–859.
- Nelson ME (2008). A mechanism for neuronal gain control by descending pathways. *Neural Comput* **6**, 242–254.
- Ohshiro T, Angelaki DE & DeAngelis GC (2011). A normalization model of multisensory integration. *Nat Neurosci* **14**, 775–782.
- Olsen SR, Bhandawat V & Wilson RI (2010). Divisive normalization in olfactory population codes. *Neuron* **66**, 287–299.
- Olsen SR, Bortone DS, Adesnik H & Scanziani M (2012). Gain control by layer six in cortical circuits of vision. *Nature* **483**, 47–52.
- Olshausen BA & Field DJ (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609.
- Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, Knight RT & Chang EF (2012). Reconstructing speech from human auditory cortex. *PLoS Biol* **10**, e1001251.
- Pi H-J, Hangya B, Kvitsiani D, Sanders JI, Huang ZJ & Kepecs A (2013). Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521–524.
- Rabinowitz NC, Willmore BDB, King AJ & Schnupp JWH (2013). Constructing noise-invariant representations of sound in the auditory pathway. *PLoS Biol* **11**, e1001710.
- Rabinowitz NC, Willmore BDB, Schnupp JWH & King AJ (2011). Contrast gain control in auditory cortex. *Neuron* **70**, 1178–1191.
- Rabinowitz NC, Willmore BDB, Schnupp JWH & King AJ (2012). Spectrotemporal contrast kernels for neurons in primary auditory cortex. *J Neurosci* **32**, 11271–11284.
- Reichardt W, Poggio T & Hausen K (1983). Figure-ground discrimination by relative movement in the visual system of the fly. Part II: Towards the neural circuitry. *Biol Cybern* **46**(Suppl.), 1–30.
- Reynolds JH & Heeger DJ (2009). The normalization model of attention. *Neuron* **61**, 168–185.
- Robinson BL & McAlpine D (2009). Gain control mechanisms in the auditory pathway. *Curr Opin Neurobiol* **19**, 402–407.
- Rothman JS, Cathala L, Steuber V & Silver RA (2009). Synaptic depression enables neuronal gain control. *Nature* **457**, 1015–1018.
- Schneider DM & Woolley SMN (2013). Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron* **79**, 141–152.

- Schwartz O & Simoncelli EP (2001). Natural signal statistics and sensory gain control. *Nat Neurosci* **4**, 819–825.
- Shapley RM & Victor JD (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J Physiol* **285**, 275–298.
- Shapley RM & Victor JD (1981). How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. *J Physiol* **318**, 161–179.
- Ulrich D (2003). Differential arithmetic of shunting inhibition for voltage and spike rate in neocortical pyramidal cells. *Eur J Neurosci* **18**, 2159–2165.
- van Atteveldt N, Murray MM, Thut G & Schroeder CE (2014). Multisensory integration: flexible use of general operations. *Neuron* **81**, 1240–1253.
- Watkins PV & Barbour DL (2008). Specialized neuronal adaptation for preserving input sensitivity. *Nat Neurosci* **11**, 1259–1261.
- Webb BS, Tinsley CJ, Barraclough NE, Parker A & Derrington AM (2003). Gain control from beyond the classical receptive field in primate primary visual cortex. *Vis Neurosci* **20**, 221–230.
- 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.
- Wen B, Wang GI, Dean I & Delgutte B (2012). Time course of dynamic range adaptation in the auditory nerve. *J Neurophysiol* **108**, 69–82.
- Willmore BDB, Bulstrode H & Tolhurst DJ (2012). Contrast normalization contributes to a biologically-plausible model of receptive-field development in primary visual cortex (V1). *Vision Res* **54**, 49–60.
- Wilson NR, Runyan CA, Wang FL & Sur M (2012). Division and subtraction by distinct cortical inhibitory networks *in vivo*. *Nature* **488**, 343–348.
- Woolley SMN, Gill PR & Theunissen FE (2006). Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *J Neurosci* **26**, 2499–2512.
- Xu X, Roby KD & Callaway EM (2010). Immunochemical characterization of inhibitory mouse cortical neurons: three chemically distinct classes of inhibitory cells. *J Comp Neurol* **518**, 389–404.
- Yin P, Fritz JB & Shamma SA (2014). Rapid spectrotemporal plasticity in primary auditory cortex during behavior. *J Neurosci* **34**, 4396–4408.
- Yost WA (1997). The cocktail party problem: forty years later. In *Binaural and Spatial Hearing in Real and Virtual Environments*, ed. Gilkey RH & Anderson TR. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 329–347.
- 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.
- Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D & Schroeder CE (2013). Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party”. *Neuron* **77**, 980–991.

Additional information

Competing interests

None declared.

Funding

Work in the authors' laboratory was funded by the Wellcome Trust. A.J.K. is supported by a Wellcome Trust Principal Research Fellowship (WT076508AIA). J.E.C. is supported by a Wellcome Trust studentship (096588/Z/11/A).

Acknowledgements

We are grateful to Nicol Harper and Ed Mann for comments on an earlier version of this review. Figure 2 is adapted from Rabinowitz *et al.* (2013), under the Creative Commons BY license, <http://creativecommons.org/licenses/by/2.0>