Comments

Comment by de Cheveigné:

Barlow (1961) suggested that the role of a sensory relay is to reduce "redundancy" defined as the number of spikes used to code incoming patterns (a slightly different usage from yours). For example if the relay has N output axons, the most likely input pattern would map to no output spikes, each of the N next most likely inputs to a single spike on a single output axon, and so on. Rare events would map to verbose patterns, their greater metabolic cost being offset by their rarity. According to Barlow such recoding might occur in stages. The net effect would be to reduce metabolic cost while providing a representation sensitive to the regularity and structure of sensory input, as well as to disruptions of this regularity. Barlow's view seems to fit the redundancy reduction you observe between IC and cortex, and in particular the probability-dependent responses found at AC but not MGB or IC. It fits also the anatomic fan-out from periphery to cortex since, according to Barlow, metabolic cost *decreases* as the number N of outputs increases, as well as low-level processing principles that minimize activity such as equalization-cancellation (Durlach 1963) or harmonic cancellation (de Cheveigné 1993). This view, if correct, has a number of sobering consequences for our quest to understand sensory processing on the basis of physiological responses. Response features at cortex are unlikely to resemble input features. Meaningful responses are likely to be abstract, sparse (i.e. one spike in one cell) and labile (if the code adapts to stimulus statistics), in other words very hard to make sense of. Conversely, easily understood stimulus-related patterns such as synchrony or tonotopy might be fortuitous remnants of input structure. See also Maass *et al.* (2003) and the comment to Elhilali *et al.* (this volume).

Barlow HB (1961) Possible principles underlying the transformations of sensory messages. In Rosenblith WA (ed) Sensory Communication. Cambridge Mass: MIT Press, 217-234.

Durlach, N. I. (1963) Equalization and cancellation theory of binaural masking-level differences. J. Acoust. Soc. Am. 35, 1206-1218.

de Cheveigné, A. (1993) Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing. J. Acoust. Soc. Am. 93, 3271-3290.

Maass, W., Natschläger, T., and Markram, H. (2003) Computation models for generic cortical microcircuits. In J. Feng (ed) *Computational Neuroscience: A Comprehensive Approach*. CRC-Press (in press).

Reply:

The ideas of Barlow had of course an important influence on our thinking as presented in this paper. The redundancy that he considered is not redundancy between neurons, but rather redundancy in the statistics of the physical stimuli, resulting from the non-uniform distribution of the natural stimuli and correlations between them. This is a crucial difference, since the two concepts of redundancy

are related in non-trivial ways, as will be discussed below. Barlow's thinking evolved over time, and his most recent summary of the redundancy reduction hypothesis (Barlow 2001) makes some important distinctions between redundancy *reduction* (creating a representation that suppresses redundancy related to the statistical structure of the natural stimuli), which he rejects, and redundancy *exploitation* (which requires explicit representation of the redundancy in the statistical structure of the natural stimuli, without suppressing this redundancy). While he still considers the economy of action potentials to be an important principle in creating useful representations of stimuli, he now interprets these representations as a way of explicitly representing the redundancy in the statistics of the natural world, with the redundancy being in fact an important property of real-world stimuli. He then goes on to suggest that under such an explicit representation of redundancy, "… the elements would also have to be active independently of each other in the *normal environment*" (my emphasis). In other words, although Barlow doesn't suggest any more that the redundancy in the statistical structure of stimuli should be eliminated in the neural representation, he claims that efficient explicit representations of the redundancy will result in informational independence between the neurons. This is in fact precisely what we see when using natural stimuli – we do not see the hallmark of Barlow's redundancy reduction, which is a sparse code with most neurons responding with high rates to only a small subset of stimuli. Rather, we observe a representation in which neurons are promiscuous, each responding with similar high rates to many different stimuli, but now the responses of different neurons are informationally independent. Thus, the redundancy reduction that we discuss is not consistent with Barlow's earlier writing, but is consistent with the most recent version of his theory.

Regarding the second part of the comment, which discusses the "sobering consequences": we mostly agree with these statements, although we feel that their pessimistic tone is somewhat too extreme. To take the points made one by one:

"Response features at cortex are unlikely to resemble input features." This is certainly true, as shown *e.g.* in Bar-Yosef *et al.* (2002) and in the chapter in this book. In specific circumstances, linearity of cortical processing (resulting in an explicit correspondence between response features and sound features) still holds (e.g. Kowalski *et al.* 1996; Schnupp *et al.* 2001), but our experience is that strong non-linear interactions appear when we use sound mixtures. However, this fact by itself need not make the decoding of cortical activity more difficult – we can still qualitatively describe the regularities in our data sets in meaningful (if more abstract) terms.

"Meaningful responses are likely to be abstract, sparse (ie one spike in one cell) and labile (if the code adapts to stimulus statistics), in other words very hard to make sense of." Again, although our results support the premise of this point, we don't necessarily accept its conclusion. The responses we record tend to have more than one spike (in contrast e.g. with responses in the barbiturate-anesthetized cat to the same stimuli), but still essentially all the information about stimulus identity is captured by a simple two-dimensional variable, using spike count and mean spike time during the stimulus (Nelken and Schnupp, in preparation). The influence of stimulus statistics is important, but can be minimized (even if not entirely eliminated, Ulanovsky *et al.* 2003) by using randomized stimulus presentation

schemes. Such experimental designs, especially if combined with long inter-stimulus intervals that minimize adaptation, make it possible to control separately the adaptation processes and the sensory responses. Although it is true that making sense of auditory cortex has already proved to be substantially harder than making sense of lower auditory stations, the advances in the last 10 years are far from being negligible.

"Conversely, easily understood stimulus-related patterns such as synchrony or tonotopy might be fortuitous remnants of input structure." We agree that the role of tonotopy or of synchrony in A1 is probably exaggerated in the literature. However, the claim as it stands is probably too extreme, as both tonotopy and synchrony shape the responses that we observe in non-trivial ways. Unpublished results demonstrate that the suppression of envelope locking to slowly fluctuating noise by low-level tones, as has been discussed in the chapter in this book, occurs only when the tone is within the sensitive area of the tonal responses. Similarly, in the bird song data, the presence of stimulus energy within the tonal response area of a neuron seems to be a necessary, although not a sufficient, condition for eliciting neural responses (Bar-Yosef *et al.* 2002). Our interpretation of the high temporal precision of some cortical spikes is also more nuanced (Fishbach *et al.* 2001; Fishbach *et al.* 2003; Nelken *et al.* 2003). As suggested by Fishbach, such well-timed spikes may represent the "edges" of auditory objects, both temporally and spectrally. Fishbach (2001) discussed a number of cases in which well-timed onset spikes in A1 follow the same rules as perceptual phenomena. One of the difficulties in A1 research is related to the presence of multiple time constants in the neural responses (as reviewed in Nelken *et al.* 2003). Some cortical responses are as precise as in the periphery, but longer time constants, on the order of 100 ms and 1-2 seconds, are also readily apparent. There is no doubt that the fastest time constants are related to peripheral processing, but all these time constants interact in non-trivial ways that seem to us crucial for elucidating the role of A1 in auditory processing.

Barlow, H. (2001) Redundancy reduction revisited. Network 12, 241-253.

Bar-Yosef, O., Rotman, Y., and Nelken, I. (2002) Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. J. Neurosci. 22, 8619-8632.

Fishbach, A., Nelken, I., and Yeshurun, Y. (2001) Auditory edge detection: a neural model for physiological and psychoacoustical responses to amplitude transients. J. Neurophysiol. 85, 2303-2323.

Fishbach, A., Yeshurun, Y., and Nelken, I. (2003) A Neural Model for Physiological Responses to Frequency and Amplitude Transitions Uncovers Topographical Order in the Auditory Cortex. J. Neurophys. (in press).

Kowalski, N., Depireux, D. A., and Shamma, S.A. (1996) Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. J. Neurophysiol. 76, 3524-3534.

Nelken, I., Fishbach, A., Las, L., Ulanovsky, N., and Farkas, D. (2003) Primary auditory cortex of cats: feature detection or something else? Biol. Cybern. (in press).

Schnupp. J. W., Mrsic-Flogel, T. D., and King, A. J. (2001) Linear processing of spatial cues in primary auditory cortex. Nature 414, 200-204.

Ulanovsky, N., Las, L., and Nelken, I. (2003) Novelty detection by auditory cortex neurons and its relation to human evoked potentials. In: Proceedings of the International