Chapter 8
From Vocal to Neural Encoding: A Transversal Investigation of Information Transmission at Long Distance in Birds

Solveig C. Mouterde

Abstract Acoustic communication in the natural world requires both emitter and receiver to adapt to the loss of information due to the transmission of sound in the environment. At the emitter's end, encoding information into propagation-resistant features may ensure its transmission on ecologically relevant distances. At the receiver's end, making sense of the degraded signal merged with added noise enables animals to produce behaviorally relevant responses. In this chapter, I use a songbird model to investigate both sides of the transmission chain, examining how individual vocal signatures are encoded, degraded and finally discriminated, with a particular emphasis on the receiver's neural encoding of degraded information. The cornerstone of this focal study is the use of naturally degraded vocalizations, combining various aspects of the challenges faced by animals performing auditory scene analysis, in order to address the complexity of real-life ecological constraints. While the individual signature of our songbird model, the zebra finch, is very resistant to propagation-induced degradation, single neurons in the avian auditory cortex have the ability to discriminate highly degraded individual vocal signatures, without prior familiarization or training. In the light of past and recent findings in birds and mammals, I delve further into significant insights uncovered by the research on neural processing and ensemble coding, and discuss the roles of perceptual plasticity and learning in the neural interface between brain and behavior.

Communication is the fabric of animal sociality. It is the foundation on which relationships between mates, rivals, or offspring develop, and population cohesion is maintained. For individuals to communicate properly, both sides of the

---

S. C. Mouterde (✉)
Equipe de Neuro-Ethologie Sensorielle ENES/Neuro-PSI, Université de Lyon/Saint-Etienne, Saint-Etienne, France

Department of Psychology, University of California, Berkeley, CA, USA

Institut de Recherche Expérimentale et Clinique (IREC), Université catholique de Louvain, Brussels, Belgium
e-mail: solveig.mouterde@uclouvain.be

T. Aubin, N. Mathevon (eds.), *Coding Strategies in Vertebrate Acoustic Communication*, Animal Signals and Communication 7,
https://doi.org/10.1007/978-3-030-39200-0_8

transmission chain must take an active part in the transfer of information. On one end, the emitter sends an acoustical message through a coded signal that can withstand some level of degradation induced by the propagation of sound through the natural environment. At the other end, the receiver(s) make(s) sense of a complex auditory stimulus (containing the potentially degraded initial signal merged with added noise) to extract the relevant information it contains, that is, to decode the signal. In this chapter, we will take a look at the whole picture, using a songbird model to investigate how the information encoded in the original signal is degraded through the environment, and how receivers deal with deciphering this altered information, both on neurological and behavioral levels.

8.1 Active Space and Encoded Information in Animal Vocalizations

In our endeavor, the notion of active space is interesting to consider because it encompasses all aspects of the communication chain. The active space of a signal is the distance from the source (emitter) over which the signal can be perceived by potential receivers (Brenowitz 1982); it is of fundamental importance in acoustic communication as it pertains to the biological relevance of propagated signals (Marler and Slabbekoorn 2004). The active space of a signal depends on its coding by the emitter, which may be the result of an adaptation to the species' lifestyle and environment; it is impacted by the acoustic constraints of the environment in which the signal propagates, which can vary in time (e.g., biotic noise, weather conditions); and it depends as well on the receivers' psychoacoustic abilities, in other words, on their capacity to perceive and decode the signal.

Within a given vocalization, the active space may also differ depending on the type of information being transmitted by the emitter. The white-browed warbler *Basileuterus leucoblepharus* provides a good illustration of this idea: the song of male warblers contains information about their species, as well as their individual identity ("individual signature," Aubin et al. 2004). Researchers found that for this bird living in the dense environment of the Brazilian tropical forest, species-specific information in the male song was encoded in a resistant acoustic feature that propagates at long distance (further than 100 m), while individually specific information degraded rapidly with propagation, restricting individual recognition to neighboring males (Mathevon et al. 2008). In this territorial species, being able to recognize the identity of neighboring males is of critical importance to mediate the male's responses to a song emitted in the vicinity; only a song perceived as being emitted by a stranger will elicit a response. Outside the territorial boundaries however, the transmission of species-specific information in the song is useful to inform potential intruders of the presence of a conspecific, but the added information about individual identity is not strictly necessary. The active space of each type of information in this species' song is thus well adapted to its ecological requirements.

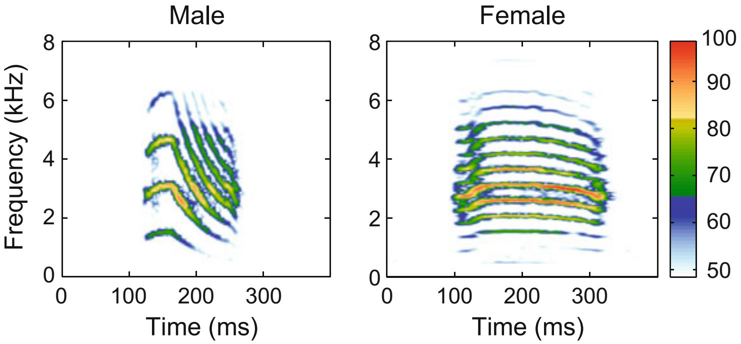

Fig. 8.1 Spectrograms of a male and female zebra finch distance call. The color scale is in relative dB as shown on the color bar, with 100 dB corresponding to the maximum amplitude observed

In other species, individual discrimination at long distance can be of critical importance, e.g., in mated pairs that need to remain within earshot when visual contact has been lost. This is the case for the zebra finch *Taeniopygia guttata*, a small gregarious songbird from subarid regions of Australia, living in large flocks in open grassy country with a scattering of trees and bushes (Zann 1996). Partners form strong pair bonds for life, and because these opportunistic breeders live in a very unpredictable environment, maintaining a strong pair bond between breeding events while living in large fission–fusion groups that are constantly on the move is of utmost importance. In this species, using a strong vocal recognition system could avoid the cost of partners losing each other. Within the zebra finch repertoire, the vocalization used for this purpose is the distance call, which was described by Richard Zann (1984) as a loud call that “will prevent members of a pair getting lost in vegetation but [it] probably serves to help them locate each other in flocks, in which contact between mates may easily be lost, especially when alarmed.” The distance call of zebra finches is a complex sound emitted by both sexes, consisting of a harmonic series modulated in frequency as well as amplitude. It is sexually dimorphic (Zann 1984; Vicario et al. 2001), the males’ fundamental frequency being higher than the females’ (typically 650–1000 Hz vs. 500–600 Hz) as well as usually being shorter and more frequency modulated (Fig. 8.1). It has been shown that the distance call bears an individual signature, and that zebra finches are capable of call-based individual recognition (Zann 1984; Vignal et al. 2004, 2008). Regarding the active space of the individual signature in these calls, it has been estimated in earlier studies, based on naturalistic observations as well as theoretical analyses, to be up to around 100 m (Zann 1996; Lohr et al. 2003).

As a common thread throughout this chapter, I will use the vocal recognition system of zebra finches as a model to study the impact of propagation-induced degradation on information content, and investigate how the birds manage to decode the degraded signals, while working within realistic biological constraints faced by this species. For this purpose, I will focus on the individual identity encoded in distance calls (that is, the individual vocal signature of the calls), and how the fine spectral and temporal acoustical structure needed for this task (Mathevon et al. 2008) is encoded in the avian auditory system.

8.2 The "Issue" of Communicating at Long Distance in a Natural Environment

When a friend talks to you from a few meters away, you will immediately recognize his/her voice. If this same friend calls you from a few hundred meters away, you might hear that it is a man or a woman calling, in any case a human being, but not necessarily that this is your friend. Acoustic signals transmitted over long distances through the atmosphere not only degrade in amplitude, but also in their spectral and temporal structure, which may alter the original information (Wiley and Richards 1982; Forrest 1994). Intensity decreases as a result of spherical spreading, as well as excess attenuation due to the nonhomogeneous environment. Changes in the spectral characteristics of the signal can be due to masking ambient noise, atmospheric absorption, and environmental filtering, whereas changes in the temporal domain may be the result of the reflection, reverberation, and scattering of sound waves (Forrest 1994). These modifications in the signal structure increase with the propagation distance, which makes the transmission of information from emitter to receiver(s) more and more difficult.

As an illustration of the difficulty of this task in the context of individual discrimination, Fig. 8.2 shows the similarity between calls of pairs of different male zebra finches as a function of propagation distance, from 2 to 256 m. The similarity, which is shown here as the spectral correlation between calls, increases with distance along with the propagation-induced signal degradation and the decrease of the signal-to-noise ratio. As an example, and to further illustrate the increasing difficulty of discriminating between individuals at long distance, the spectrograms of the same calls from two different males are shown for various propagation distances (Fig. 8.2). Note however that this propagation-induced degradation of the signal is not solely a hindrance in animal communication, as receivers can use it to locate the emitter and estimate the propagation distance (a process called "ranging"; Naguib 1995; Fotheringham 1997; Holland 1998).

At the receiver's end, the challenge is to interpret this sensory information in the auditory system and display behavioral responses that are adapted to the situation and context. This ability to extract relevant information from a complex environment and to interpret it in terms of behaviorally relevant objects, otherwise known as scene analysis, is universal to species and sensory modalities (Lewicki et al. 2014). This process is especially complex in the auditory context where sounds from various sources and localizations, potentially degraded by transmission through the environment, combine to form a single acoustical signal. While a number of studies have highlighted the proficiency with which animals deal with this difficult task of discriminating what is relevant from what is not in degraded auditory signals

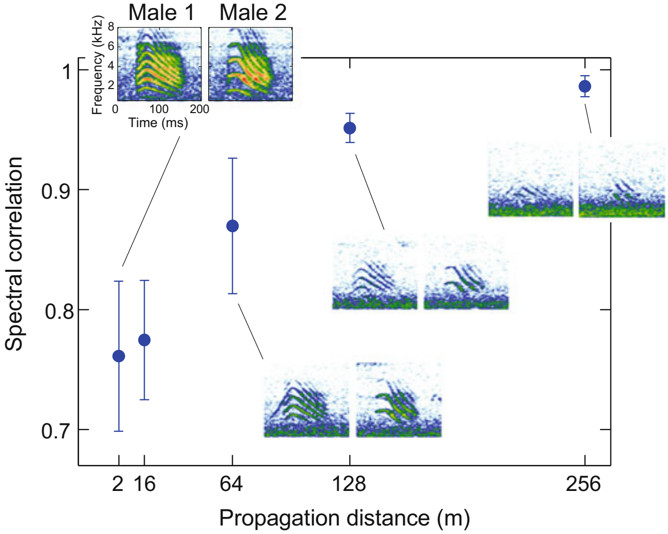

Fig. 8.2 Spectral correlations (i.e., similarity) between the distance calls of 17 pairs of male zebra finches, as a function of propagation distance. Mean correlations are represented for each distance, and error bars correspond to the standard deviation. Correlations were calculated between the mean frequency spectra of each male. As an example, the spectrograms of the same calls from two different males are shown for various distances (reproduced with permission from Mouterde et al. 2014a)

(Schnitzler and Flieger 1983; Aubin and Jouventin 2002; von der Emde 2004; Appeltants et al. 2005), understanding the underlying biological mechanisms involved in auditory scene analysis has proved particularly challenging for scientists (Shamma et al. 2011), especially on the neurophysiological level. In order to discriminate individual identity from vocalizations propagated at different distances, neurons must show a certain degree of perceptual invariance, that is, some stability in their responses to stimuli showing large variations in their acoustic parameters (Bregman 1993). A number of studies dealing with auditory scene analysis and individual discrimination or recognition in birds have revealed important insights on specific aspects of this question, such as invariance to intensity or background noise (Billimoria et al. 2008; Moore et al. 2013; Schneider and Woolley 2013), or individual recognition of undegraded signals (Chew et al. 1996; Gentner 2004). However, concern has been recently raised that studies on scene analysis do not address the complexity of the problems that need to be solved in natural settings; researchers were thus encouraged to focus more on the real-life issues faced by animals in their natural environment (Lewicki et al. 2014). Furthermore, since the auditory system of animals has evolved to process behaviorally relevant natural sounds (Mizrahi et al. 2014), using natural stimuli to investigate auditory scene analysis would be the best way to understand the neural computations at play (Theunissen and Elie 2014).

In the study presented as a common theme in this chapter, naturally degraded vocalizations were used in order to reflect ecologically relevant stimuli and behaviors for our songbird model. The degraded calls shown in Fig. 8.2 were recorded in a natural environment, on a flat field and in low wind conditions, at a height of 1.30 m and at propagation distances ranging from 2 to 256 m (for more details on the recordings, see Mouterde et al. (2014b)). In these calls, the reduction of the signal-to-noise ratio is due to the combined effects of attenuation, sound degradation, and ambient noise. In order to explore the real-life issue of partner acoustic recognition in our zebra finch model, we used these naturally degraded calls to investigate individual discrimination in different facets of the communication chain, from emitter to receiver.

8.3 The Coding of Individual Vocal Signatures in Propagated Calls

In the natural environment, communication distance is affected by a number of factors pertaining to the signal's emitter. An obvious one is the loudness of vocalizations, which may have evolved to serve a territorial function in the case of songs for example (Brenowitz 1982), and which can in any case vary due to proximal causes such as noise (Cynx et al. 1998; Brumm and Todt 2002). Active space can be enhanced (or conserved in adverse environmental conditions) by repeating the signal and thus increasing the redundancy of encoded information, or emitting the signal at a certain localization or time (Brumm and Naguib 2009). While these findings shed light on the emitters' capacity to adapt to various transmission conditions, investigating how the information is encoded in the spectro-temporal features of the vocalization itself, and how it is affected by long-distance propagation, is critical to understand how the receivers solve the issue of interpreting degraded signals at the other side of the communication chain. The first studies investigating the actual information content of propagated calls mostly dealt with species-specific information, providing significant insights into the link between birds' vocalizations and their habitat. Indeed, the coding of information in acoustic signals is generally based on features that are resistant to propagation, depending on the environmental acoustics that are characteristic of the habitat (Morton 1975; Mathevon and Aubin 1997). Generalities can be drawn: the most important acoustic features for species discrimination are usually found in a frequency bandwidth that is less degraded through propagation (Brenowitz 1982), and within this bandwidth, frequency modulation can mediate the discriminability of the information (Brémond and Aubin 1990; Mathevon et al. 1997). These observations form the basis for further investigations into the long-range degradation of finer grained information such as individual identity. In this respect, research has been rather scarce. Two previous studies using birds have been published on the subject; in the white-browed warbler, mentioned earlier in this chapter, the individual signature is encoded in songs in a succession of pure tones decreasing in frequency (Aubin et al. 2004; Mathevon et al. 2008). Conversely, in the male corncrake *Crex crex*, individuality is coded in propagated calls in the inter-pulse duration (Rek and Osiejuk 2011). In these examples, individual signatures are encoded in either the spectral or the temporal domain, with pure tones or pulses.

Let us now consider the propagation-induced information loss in complex sounds, i.e., in vocalizations displaying wide frequency bandwidths together with amplitude and frequency modulations. The conundrum in this case is to calculate the appropriate parameters for the acoustic analysis: in calls propagated at long distance, it is not possible to describe precisely the spectro-temporal features such as parameters following closely the fundamental frequency and its variation in time. These parameters, frequently used in studies investigating individual signatures (Guyomarc'h et al. 1998; Naguib et al. 2001; Vignal et al. 2008), are impossible to calculate here with reasonable precision, with the fundamental frequency progressively disappearing within the background noise, and the low resolution of the temporal characteristics hindering the calculation of even basic measurements such as the calls' duration. In the first step of our focal study, this challenge was tackled by performing a sound transmission experiment using zebra finch calls, with the aim of characterizing and quantifying the degradation of the calls' individual signature during propagation. To this end, 16 exemplars of distance calls from each of 16 female and 16 male individuals (that is, 512 calls) were recorded at five propagation distances, from 2 to 256 m. The calls were then characterized using two distinct sets of acoustic parameters. The first set, which will be called "envelope parameters", described separately the spectral envelope (amplitude in the spectral domain) and the temporal envelope (amplitude in the temporal domain) of each call. The second set used the complete spectrogram, in an attempt to circumvent the use of subjective assumptions on the nature of information-bearing acoustical features in the calls. A principal components analysis was used to reduce the high dimensionality to a manageable set of parameters describing the variability of the calls, subsequently called the "spectrogram principal component" or "SPC" parameters. To test for the presence of an individual acoustic signature in the propagated calls, discriminant function analyses were performed separately for each sex and each propagation distance (Mouterde et al. 2014b). These analyses yielded percentages of correct classification of calls (PCC), which quantified the discriminability of different individuals within each sex and for each distance; a percentage value over the chance level (1 over 16 individuals) shows the existence of individualized acoustic features in the calls.

We found that the individual vocal signature encoded in the distance calls of zebra finches is remarkably resistant to propagation-induced sound degradation. Figure 8.3 shows the percentages of correct classification, that is, the individual discriminability, as a function of propagation distance, for each set of parameters and each sex.

While the individual discriminability of calls logically decreases with distance, it is still well over the chance level at 128 m for both sets of parameters, and up to 256 m for the SPC parameters. SPC parameters yield higher discriminability values than envelope parameters, especially at longer distances, which can be explained by the fact that while the envelope parameters take spectral and temporal features into

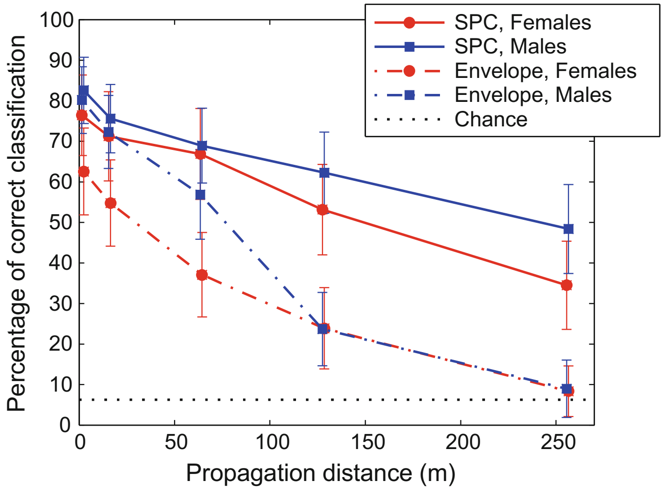

Fig. 8.3 Mean percentages of correct classification (i.e., individual discriminability) obtained for each propagation distance and both sexes. SPC parameters are represented as solid lines and envelope parameters as dash-dot lines. Standard deviations are indicated. The chance level, corresponding to 6.25% of correct classification, is shown as a horizontal dotted line (reproduced from Mouterde et al. 2014b)

account separately, SPC parameters are extracted from the full spectrogram, thus also describing the joint spectro-temporal variations in the calls such as up- or down-sweeps (Mouterde et al. 2014b). We also found that at short distances, both sets of parameters yield similar discrimination performances for males (the performance of the envelope parameters decreasing rapidly at 64 and 128 m), while for females the envelope parameters lead to noticeably lower discriminability values. Hence, the added information available in the SPC parameters is not indispensable for discriminating between males at short distances, the information extracted separately from the spectral and temporal domains (envelope parameters) being sufficient to yield equally high discrimination performances. This difference in coding properties is well illustrated in Fig. 8.4, which shows the first three discriminant functions (DFs) calculated in the analysis for the SPC parameters, represented in the spectrographic space at short, medium, and long distance. In other words, Fig. 8.4 is a description of the most important features in the spectrogram that can be used to discriminate between individuals at various distances: red bands show features whose presence is important for individual discrimination, while blue bands show features whose absence is important. One can observe that while the available frequency bandwidth shrinks with distance, fine spectral cues (as seen on the stacks of red and blue narrow bands) are useful at all distances, and for both sexes. Interestingly, at short and medium distance, the DFs pertaining to male calls do not look at all like the spectrographic representation of an average male call, the faster and more predominant down-sweep component in particular (seen on Fig. 8.1) being absent for short and medium distances and only appearing at 128 m (not shown on Fig. 8.4) and 256 m (DF2 and DF3, Fig. 8.4).

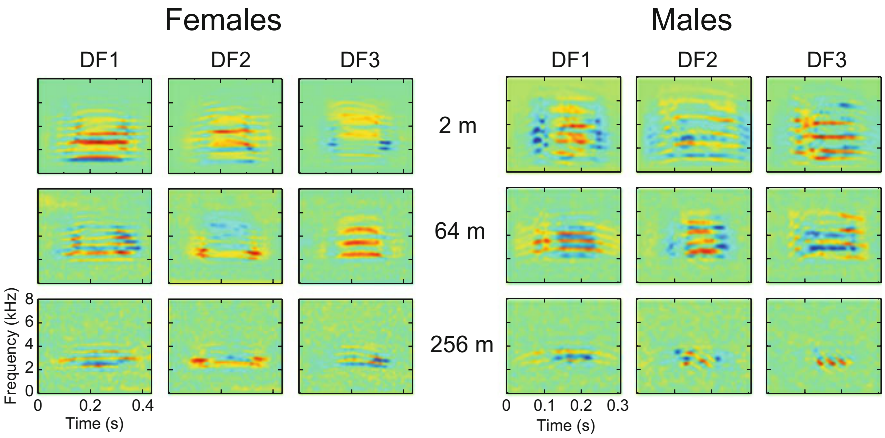

Fig. 8.4 Representation of the discriminant functions (DFs) projected into the spectrographic space for the SPC parameters. The first three DFs, obtained from the SPC parameters, are shown for females (left) and males (right). Each row indicates the propagation distance used to perform the discriminant function analysis. Positive frequency bands are shown in red and negative bands in blue. This representation describes the most important features in the spectrogram that can be used to discriminate between individuals at various distances (adapted with permission from Mouterde et al. 2014b)

This corroborates our previous conclusions, in the sense that while information about frequency modulation is not necessary to discriminate between males at close range, it becomes critical at long distance in order to maintain individual discriminability above the chance level. Thus, while it was already known that the downsweep component of male calls is highly individualized (Vignal et al. 2008), this study shows that the same level of individualization can be found at close range in male calls using parameters describing only its energy spectrum and its temporal variation in amplitude. This redundancy in the coding of the individual signature, clearly shown here for the male calls, is also found for the female calls of zebra finches, albeit to a lesser extent (Mouterde et al. 2014b); indeed, in females, the information about spectro-temporal modulations is important at all distances to obtain higher discrimination performances.

In a real-life situation, redundant coding of information within a single short call can make all the difference in terms of communicating efficiently in ever-changing conditions. Let me be clear: the furthest propagation distance (256 m) advertised here as still enabling the transmission of individualized information in the distance calls of zebra finches is the result of a transmission experiment performed on a specific day, with specific weather conditions and background noise. While testing at longer distances might have pushed the boundary of theoretical active space a little further, tougher recording conditions (wind, biotic noise) could have drastically reduced it (Brumm and Naguib 2009). In this respect, redundantly coded information in vocalizations may help maintain efficient communication at shorter distances despite the varying conditions in the environment in which communication takes place.

We will now turn to the other side of the communication chain to investigate how receivers optimize the gained information and in particular, which neurophysiological processes are at play to make sense of this degraded acoustic information.

8.4 Signal Decoding: Investigating the Neural Substrate for Auditory Scene Analysis

8.4.1 Increasing Selectivity for Behaviorally Relevant Features Along the Auditory Pathway

In order for a bird to appropriately respond to information encoded in a naturally propagated vocalization, it must have the neural ability to make sense of this auditory input. This process involves discriminating the relevant features in perceived sounds, that is, showing a higher selectivity for the type of information that is relevant to the animal. Indeed, as auditory information from the sensory system ascends the avian auditory pathway through the midbrain, thalamus and into the auditory cortex, itself divided into a primary (field L complex, a region that is analogous to the primary auditory cortex of mammals) and a secondary region (caudal mesopallium or CM; caudomedial nidopallium or NCM), neurons show an increasing selectivity in their responses for natural sounds (Theunissen et al. 2004; Woolley 2006). In particular, high selectivity for conspecific songs has been found in both primary and secondary regions of the auditory cortex (Grace et al. 2003; Hsu et al. 2004; Theunissen and Shaevitz 2006; Hauber et al. 2007). These findings support a hierarchical view of auditory processing, in which neurons in the lower pathway tune to simple low-level acoustic features and send their combined inputs to the forebrain neurons, which in turn detect more complex natural sound features (Woolley et al. 2009; Meliza et al. 2010; Theunissen and Elie 2014). This hierarchical sensory processing enables the representation of sounds as auditory objects (i.e., stable perceptual units), reflecting the system's transformation of an acoustic waveform from low-level sensory representation into perceptual representations that are behaviorally relevant to the organism (Christison-Lagay et al. 2015). A number of studies suggest that further hierarchical processing occurs within the auditory cortex (Meliza and Margoliash 2012), with secondary auditory areas encoding stimulus surprise (Gill et al. 2008) or playing a major role in learned auditory discriminations (Gentner 2004; Pinaud and Terleph 2008; Jeanne et al. 2011).

In the context of auditory scene analysis, signal decoding also implies extracting the biologically relevant information content from the noise in the perceived sound, whether the decrease in signal-to-noise ratio (SNR) stems from the noise level itself or the reduction of signal intensity due to propagation. This process starts in the peripheral auditory system, in which a selective frequency tuning process breaks complex sounds into their frequency components (Fay and Popper 2000). This tonotopic organization can improve the SNR for the representation and detection of acoustic signals by reducing the masking interference from background noise when signal and noise are in different frequency bandwidths (Klump 1996; Fay and Popper 2000). Further along the ascending pathway, neurons in Field L have been shown to adjust their spectral tuning properties to the characteristics of background noise, enhancing their frequency selectivity as a result (Nieder and Klump 1999). This dynamic sharpening of auditory filters in the auditory forebrain of birds echoes the idea, also investigated in mammals, that higher-level neural processing enables an increased tolerance for noise (Rabinowitz et al. 2013). Indeed, two electrophysiology-based studies on zebra finches found noise-invariant coding in neurons in NCM responding to familiar and unfamiliar songs from conspecifics (Moore et al. 2013; Schneider and Woolley 2013). Another study described a population of field L neurons able to discriminate between bird songs while being invariant to intensity, showing a link between invariance and discrimination performance (Billimoria et al. 2008). This increased tolerance in the auditory forebrain for variations in the signal's characteristics would indeed be an important factor of success in the task of discriminating between individuals at a distance in a natural environment.

8.4.2 Discrimination of Naturally Degraded Individual Vocal Signatures in Single Cortical Neurons

In these studies above, the overall acoustical quality of the vocalizations used as stimuli was preserved, and thus the question of how the songbird brain deals with the impact of propagation-induced degradations as naturally experienced by the animals when communicating at long range was yet to be investigated. Using the previously described database of propagated distance calls, we tested zebra finches' brains with some rather tough questions: are there single auditory neurons that are able to discriminate such fine-detailed information as individual identity in short vocalizations like distance calls, using not simply individual calls but multiple call exemplars from each individual, after they have been naturally degraded through propagation? Is such a discrimination capacity possible spontaneously, without prior experience of long-range propagation or prior knowledge of the individuals used as stimuli, and with little help from higher order neural processes such as attentional mechanisms? To answer these questions, extracellular electrophysiological recordings were performed on anaesthetized zebra finches, spanning the whole auditory forebrain, to test whether neurons from these aviary-reared subjects could discriminate individual identity in unfamiliar degraded distance calls of males or females (stimuli from each sex being tested separately, on different recording sites). From this multiunit data we identified single auditory units and analyzed their responses (Mouterde et al. 2017). We found neurons showing high discrimination of the identity of unfamiliar vocalizing birds, at short as well as long distances: these neurons are able to code distinct individuals differently, and maintain some aspects of this code across all tested distances, from 2 to 256 m (Fig. 8.5). As an example, a

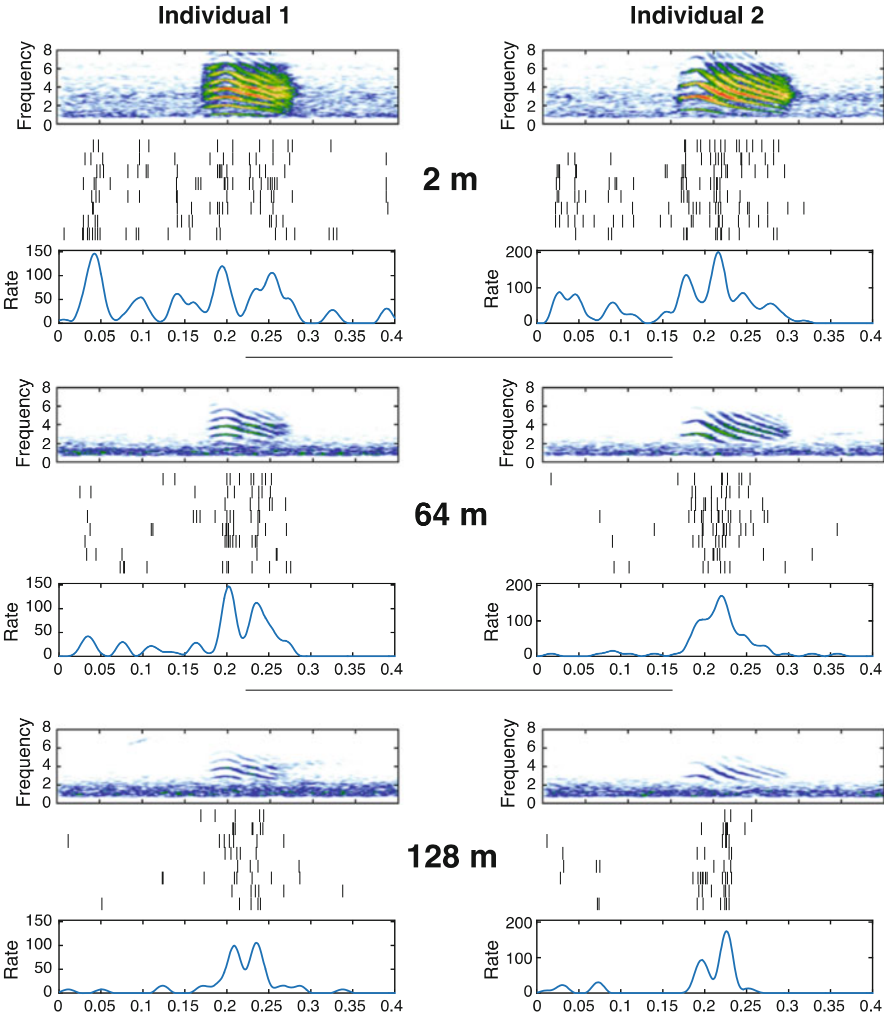

Fig. 8.5 Comparison of the responses of a unit highly discriminative for the signature of individual vocalizers to the call of two different males, at three tested distances. For each propagation distance, the spectrogram of the stimulus call is shown on top, followed by the spike trains for the eight recorded presentations, and by the peristimulus time histogram averaging these eight presentations. Sound frequency is given in kilohertz and rate in spikes per second (adapted with permission from Mouterde et al. 2017)

high-performance neuron discriminating between multiple calls from four different males showed a percentage of correct classification (PCC) over the chance level of 66.8% at 2 m, and 34.4% at 256 m (the chance level being 25% in this case). This impressive ability would require a high degree of invariance to decreasing signal-to-noise ratio and/or sound intensity, as well as the ability to extract parameters encoding individual identity in calls from previously unknown vocalizers.

In order to disentangle the overall effect of sound propagation through natural environment from the mere effect of intensity decrease, we compared the units' responses to naturally degraded calls to their responses to synthetic calls that mimicked the intensity decrease of the naturally degraded calls, but with the same high SNR as the one observed for the calls recorded at 2 m. We found that these units are remarkably invariant to the signal degradation and to large decreases in SNR induced by long-range propagation, the decline in discrimination performance being mostly a result of the intensity decrease (Mouterde et al. 2017). Interestingly, we did not find that the neural substrate for the discrimination of individuals in degraded calls was linked to specific traditional auditory areas (field L, NCM etc.), but we did find a distinctive spatial distribution of these neurons, the units most discriminative of individual identity being mostly found in superficial as well as deep regions of the auditory cortex.

Looking further into the coding properties of these discriminant neurons, we examined the spikes (or action potentials) which fired in response to the propagated calls, and calculated for each unit the length of the encoding time window yielding the highest discrimination performance for individual identity; the encoding time window is defined as the duration of the spike train assumed to correspond to a single symbol in the neural code (Theunissen and Miller 1995). This analysis also enabled us to investigate different temporal resolutions for the neural code depending on the length of the time window, exploring at one end the mean firing rate over the stimulus (rate code), and on the other end the fine temporal structure of the spike train, that is, the actual patterns of spikes within the encoding time window (temporal code). We found that while the average spike rate during the whole stimulus presentation was related to the discrimination performance, the precise timing of the spike patterns within a shorter encoding time window provided more information regarding the coding of individual identity in degraded calls. This observation is in line with a number of studies showing that neurons use a spike timing strategy to encode vocalizations (Narayan et al. 2006; Huetz et al. 2011; Gaucher et al. 2013; Lopes-dos-Santos et al. 2015). Furthermore, sensory neurons have been shown to encode different stimulus attributes on different temporal scales, resulting in a temporal multiplexing of information (Panzeri et al. 2010). We investigated how the units coded for propagation distance as well as individual identity, and indeed found units coding for both types of information, using respectively longer (~30 ms) and shorter (~10 ms) encoding time windows (Mouterde et al. 2017). This multiplexing ability enhances the encoding capacity of neural responses, enabling disambiguation of stimuli that cannot be discriminated at a single timescale and making sensory representations more stable in regard to variability. Temporal multiplexing is thus acknowledged as a strategy used by the brain to create information-rich and stable representations of noisy and variable environments (Panzeri et al. 2010), and could be a key factor in efficiently decoding the individual signature in propagated vocalizations.

It is interesting to compare the performance level of the most discriminant neurons to the performance obtained in the acoustic analysis previously described. For males for example (and using SPC parameters, see Sect. 8.3), the mean PCC over the chance level is 76.4% at 2 m and 42.1% at 256 m following the acoustic analysis, compared to respectively 66.8% and 34.4% for the most discriminant neuron tested with male calls. Although these values correspond to two different assessments of discriminability (statistical discrimination from the calls' acoustical structure versus between-calls discrimination ability by one particular neuron), this comparison implies that some neurons are able to take the best advantage of the information still remaining after long-distance propagation. Indeed, these observations highlight the fact that neurons may be able to retrieve a comparable, if somewhat lower amount of information at long distance than what was theoretically calculated using acoustical features optimized for that distance. These results show that complex processing of higher order acoustic features occurs at the level of the single neuron.

8.4.3 Ensemble Coding: The Auditory System as a Neural Network

We will now take a step back and consider the brain, processing information through a network of interconnected neurons, each of which carrying complementary information about the auditory stimuli. While single neurons can encode information on different timescales, ensembles of neurons can also process information on different spatial scales, their synchronization properties potentially generating an additional layer of sensory processing (Quian Quiroga and Panzeri 2009; Panzeri et al. 2010). Investigating the population code, defined as a neural representation in which information is conveyed by relative amounts of activity across multiple elements of an array (Covey 2000), has thus become the focus of a number of neurophysiological studies. Feng and Ratnam (2000) stated that an important and unresolved issue was to determine at which level of complexity certain features are represented in single neurons while others are represented across groups of synchronized neurons. It has been found since that information processing by single neurons can reach a level of complexity that had not been shown before in the context of auditory scene analysis (Mouterde et al. 2017). Of course, since each individual neuron receives excitatory and inhibitory inputs from a population of neurons that are lower in the ascending pathway, the analysis of single neurons in the auditory cortex is in itself a "readout" of a population code, each neuron's response being a computation of all these convergent inputs (Covey 2000). Taking into account the coordinated activity of a population of cortical neurons, however, could provide access to higher-order levels of sensory information and enhance the information-carrying capability of the auditory cortex (Woolley 2006; Kayser et al. 2009). Depending on the tuning curves of the neural population involved, a population code would either increase the range of stimulus features encoded by that population, or lead to a better discrimination of stimulus features (Panzeri et al. 2015). Previous studies on mammals suggested that ensemble coding could be critical to various auditory processes, such as auditory stream formation (defined as a series of sounds that is perceived by the listener as a coherent entity; Fishman et al. 2001; Shamma et al. 2011), sound source localization (Fitzpatrick et al. 1997; Miller and Recanzone 2009) or sound level encoding (Dean et al. 2005). Another interesting observation is that the brain encodes sensory information using a small number of active neurons at any given point in time, this strategy being referred to as “sparse coding” (Olshausen and Field 2004). Regarding auditory scene analysis, a spatially distributed population code, much like temporal multiplexing, can stabilize the sensory representation to external or internal noise (Kayser et al. 2009).

In birds, the representation of complex sounds by synchronized activation of groups of neurons has been observed in the song system (Margoliash et al. 1994) and the encoding of unexpected auditory events (Beckers and Gahr 2012). Further research is needed to analyze ensemble codes in the field of sensory processing; in particular, it remains to be determined how further gains in the neural discrimination of identity in naturally propagated calls could be achieved by examining the population code. I hypothesize that we could estimate the stimulus intensity from the ensemble responses of cortical neurons; as explained earlier, the most discriminant units for individual identity in degraded calls that we found were mostly affected by intensity decrease while being highly invariant to large decreases in SNR and degradations of the signals’ spectro-temporal features (Mouterde et al. 2017). Investigating the neural population code could improve the sound level coding accuracy (Dean et al. 2005). In recent years, a number of new analysis techniques for information processing in neural networks have been developed, which would be of interest to researchers aiming to study these questions further (Mahmud and Vassanelli 2016). In addition to spatial population coding (including accounting for the various encoding timescales of the single neurons encompassed in the recorded population), imaging techniques such as functional magnetic resonance imaging have proved useful (Boumans et al. 2008), as well as the analysis of mass signals such as local field potentials (LFPs). LFPs are extracellularly recorded potentials with frequencies of up to ~500 Hz that reflect the slow fluctuations of neural activity of local cortical networks, thus providing an aggregate measure of local sub-threshold integrative processes and network state changes that are not reflected in spike trains (Kayser et al. 2009; Einevoll et al. 2013). It has been found that combining the analyses of simultaneously recorded LFPs and spikes provided more information about cortical activity (Quian Quiroga and Panzeri 2009; Einevoll et al. 2013). In particular, the nested combination of spike patterns with the phase of low-frequency network rhythms was highlighted as an especially promising code, as it was found to be most informative and provided robustness to noise (Kayser et al. 2009). Slow cortical rhythms thus help stabilize neural codes to the detrimental effects of sensory noise, and taking into account such nested combinations of codes in the study of auditory scene analyses might majorly increase our understanding of the processes involved in the discrimination of individual signatures in naturally degraded vocalizations.

As tempting as it may be, however, to think that we are now close to finding the key to how the brain as a neural network decrypts information encoded in sounds, and to predicting reliably the subsequent behavioral responses, some important issues need to be addressed. First, given the sheer size of the auditory system’s network of interconnected neurons, most decoding algorithms may fail to decode stimuli owing to a high-dimensional response space or the use of incorrect assumptions about the neuronal code actually used (Quian Quiroga and Panzeri 2009). While recent studies have focused on methods for dimensionality reduction (Cunningham and Yu 2014), interesting findings suggest that a small-dimensional subspace of the experimentally measured activity may be sufficient to explain the population dynamics underlying sensory processing (Bathellier et al. 2012; Panzeri et al. 2015). Second, neural responses are sensitive to context, which constitutes a substantial source of variability. This may relate to sounds that have occurred at specific times prior to the stimulus under consideration (Covey 2000; Mizrahi et al. 2014); an illustration of this sensitivity to context is the fact that we are better at discriminating frequency relationships (e.g., recognizing a melody regardless of the key in which it is played) than we are at making absolute frequency judgments, (e.g., having absolute pitch). Context also relates to the internal state of the brain, which may or may not be mediated by the animal's environment. Cortical states can vary along a continuum of synchronized and desynchronized states. A synchronized state (commonly observed during slow-wave sleep and anesthesia) is characterized by slow fluctuations between intrinsically generated up and down states, while in a desynchronized state (commonly observed during sensory processing in awake animals), activity is strongly modulated by sensory inputs (Pachitariu et al. 2015). It has been found in mammals that variations in brain state accounted for a significant amount of the variability in population activity, the cortical state modulating the selectivity and reliability of spike patterns and impacting the fidelity of the population code as a result (Curto et al. 2009; Pachitariu et al. 2015). This cortical state can vary with behavioral and cognitive states, and may play a key role in shaping the responses of the awake cortex to sensory stimuli (Curto et al. 2009). In anaesthetized animals, the internal synchronous activity of neural networks has been found to be involved in the neural processing of context-dependent auditory information in the secondary auditory areas of zebra finches (Beckers and Gahr 2012). Since the behavioral state has a major impact on the neural interpretation of auditory stimuli, and since a neuronal code yielding high information values might not be biologically relevant in the sense that neural systems might not be capable of exploiting all of this information (Quian Quiroga and Panzeri 2009; Nagel et al. 2011), it is critical to understand how neural activity in sensory cortices relates to perception, so as to bridge the gap between neural processing and behavior. This is how we will come closer to understanding the neural interface between animals and their environment.

8.5 In Search of the Neural Interface Between Brain and Behavior

8.5.1 The Role of Conscious Perception on Neural Processing

We will now turn toward the role of attentional and learning mechanisms in the neural encoding of auditory information. In the zebra finch study described earlier (see Sect. 8.4.2), the subjects were anaesthetized with urethane, which, as we saw above, would induce a different brain state than that of awake animals. What, then, would be the potential benefits of the attentional mechanisms found in awake animals in terms of discrimination of auditory objects? A number of studies have compared neural activity in both awake and anaesthetized songbirds in response to vocalizations, and produced controversial results. Studies examining the impact of urethane anesthesia on the midbrain (Schumacher et al. 2011), primary (Narayan et al. 2006) and secondary auditory cortex (Meliza et al. 2010) of zebra finches have found that while some parameters (such as neural excitability, spike precision or inter-trial correlation) differed between anaesthetized and awake states, anesthesia did not affect the discrimination of songs or selectivity for song motifs. In contrast, other studies using songbirds found important changes in neuronal preferences between both states (Capsius and Leppelsack 1996; Schmidt and Konishi 1998; Karino et al. 2016). In an effort to address this apparent discrepancy, Karino et al. (2016) argued that anesthesia might modify the balance of neuronal preference between behaviorally significant auditory signals and nonsignificant sounds; vigilance and attention may thus be required in order to produce the appropriate behavioral responses toward particularly meaningful vocalizations. Interestingly, in a study investigating neurons in the secondary auditory cortex of anaesthetized zebra finches that responded preferentially to unexpected sounds (on the basis of recent stimulus history), Beckers and Gahr (2012) found an internally synchronized neural network that may reflect an early-stage process involved in the involuntary capturing of attention. In any case, attentional processes might indeed be useful in the task of extracting individual signatures from degraded and potentially noisy signals, which is faced by animals in their natural environment. In humans, Christison-Lagay et al. (2015) found that selective attention acts by enhancing the cortical representations of the attended sound streams, the population-level activity in the presence of competing sound streams reflecting primarily the activity in response to the attended stream in isolation compared to that of the ignored stream.

These observations spark the debate on the relative involvement of ascending vs. centrifugal pathways in the perception of behaviorally relevant auditory objects. On one hand, the neural representation of sounds is transformed systematically along the ascending auditory pathway, leading to integration in time, frequency, and space domains and allowing the extraction of behaviorally relevant features (Feng and Ratnam 2000). This “bottom-up” process is driven predictably by patterns of acoustic variations in the stimulus (Gentner 2004). On the other hand, this information flow is subject to dynamic modulation by a descending auditory pathway that runs from the cortex to the cochlea and which is largely independent of any particular acoustic feature (Feng and Ratnam 2000; Gentner 2004). These “top-down” mechanisms enhance the response of some neurons while decreasing the response of others (Jen et al. 1998), and may thus form the basis for attention, motivation, and/or reward mechanisms, enabling animals to acquire information about one or two auditory objects while ignoring others. The role of this centrifugal pathway is particularly useful for the extraction of signals in difficult listening conditions (Feng and Ratnam 2000). In this light, the interpretation of an auditory scene consists of a selection of particular auditory objects through a combination of bottom-up object salience and top-down attention, filtered by experience and expectation (Middlebrooks and Simon 2017).

While we have found that bottom-up processes can discriminate very fine-detailed information such as information about individual identity and propagation distance in highly degraded calls (Mouterde et al. 2017), the next step is to investigate how top-down processes, in the form of conscious perception, can further improve sensory processing. Can a bird learn to better discriminate degraded information in these signals over time?

8.5.2 Perceptual Learning: Pushing the Limits of Discrimination

Perceptual learning, a manifestation of experience-dependent plasticity in the sensory systems, occurs at the developmental stage as well as throughout adult life (Seitz and Dinse 2007; Dahmen and King 2007). It involves improved sensitivity independent of cognitive, motor or other, non-perceptual factors (Gold and Watanabe 2010). Interestingly, while attention and behavioral reinforcement mediate the selection and learning of only the information that is deemed to be of importance, perceptual learning has been found to occur not only under training conditions but also in situations that lack attention and reinforcement (Seitz and Dinse 2007). As an example, unreinforced learning of song discrimination has been shown to occur rapidly in zebra finches, after 3 hours of passive song exposure (Stripling et al. 2003). Seitz and Dinse (2007) have thus suggested that in order for learning to occur, sensory stimulation needs to be sufficient to drive the neural system past the point of a learning threshold. The added “layer” of attention and conscious perception however seems necessary for memory consolidation (Gilbert et al. 2001). The study of perceptual plasticity at the receiver’s level has been generally overlooked in auditory scene analysis, comparatively to the extensive research that has been conducted on the vocal plasticity of emitters in the context of difficult communication conditions (Pohl et al. 2012; Slabbekoorn 2013). While perceptual learning may increase the discrimination of specific acoustic features, how efficient can it be when approaching the limit of sensory perception? In other words, at which point will the inherent limits of the auditory system hinder any potential improvements due to perceptual learning?

As a last step in our focal study on the discrimination of degraded vocal signatures in zebra finches’ calls, we investigated the role of experience in the discrimination performance of highly degraded calls. Adult female zebra finches were used in forced-choice conditioning experiments, in which the results of two different protocols were compared. In both protocols, the females were asked to discriminate between the propagated calls of two unfamiliar males. In the first protocol, the subjects learned to discriminate between a number of calls of two males recorded at short range before being systematically challenged with the calls of the same males recorded at longer distances. In the second protocol, the subjects were challenged daily with a different pair of males and with a randomly selected propagation distance, and thus did not have the possibility to learn from their previous experience in the task. We found that the subjects were able to discriminate between the degraded calls of male zebra finches at up to 128 m without training, and up to 256 m with training. This augmented performance in the training condition suggests that the subjects enhanced their discrimination abilities through perceptual learning, even in a very difficult listening situation (Mouterde et al. 2014a).

On a physiological level, while it is a challenge to identify the changes in neuronal response properties that are causally related to the perceptual improvements, evidence for learning-related changes in the auditory cortex have accumulated, with the learning process resulting in an expanded representation of the trained features (Recanzone et al. 1993). Indeed, devoting more neurons to processing those aspects of a sensory stimulus that have particular behavioral relevance may seem like a plausible way of improving perceptual acuity (Dahmen and King 2007). In songbirds, the modification of neurons' encoding properties following experience-dependent perceptual plasticity was found in the secondary auditory regions, CMM (Gentner and Margoliash 2003), NCM (Pinaud and Terleph 2008), and in a lesser extent, CLM (Jeanne et al. 2011)—the latter authors suggesting that CLM and CMM are a part of a functional hierarchical neural circuit. These regions contribute to the recognition of familiar songs, and the learning of novel song discriminations (Gentner et al. 2004). Another important issue is to disentangle the effects of bottom-up and top-down processes in perceptual learning (Gentner 2004). While some studies supported the idea of a top-down control of perceptual learning, showing its strong interaction with attention (Ahissar and Hochstein 1993; Gilbert et al. 2001), other studies showed that perceptual learning can occur not only under training conditions but also in situations of unattended and passive sensory stimulation (Stripling et al. 2003; Seitz and Dinse 2007). Overall, it is probable that the plasticity in adult cortical networks that accompanies perceptual learning is shaped by an interaction between bottom-up sensory inputs, neuromodulator release, and task-specific top-down inputs (Gentner 2004; Polley et al. 2006; Dahmen and King 2007). Training facilitates these processes by strengthening both bottom-up sensory encoding and top-down modulation of the auditory cortex (Caras and Sanes 2017).

8.5.3 A Neural Substrate Readily Available for the Discrimination of Fine-Detailed Information

While we have found that zebra finch females could not discriminate between the naturally degraded vocal identities of males at 256 m without training, some single neurons in naive birds had the capacity to perform this task. Keeping in mind that both studies involved different subjects (but the same auditory stimuli), it seems that for a comparable level of training, this discrimination ability found at the neural level is not shown at the behavioral level. This finding is interesting in regards to previous studies that have shown in a number of sensory systems that behavioral discrimination performance tends to mirror the neuronal discrimination of the most selective single cells (Britten et al. 1992; Romo and Salinas 2003; Wang et al. 2007). In the auditory system, behavioral performance for song discrimination matched the neural performance of the best neurons using a spike-timing code, the activity of which might potentially reflect the computations performed by the entire network (Wang et al. 2007; Narayan et al. 2007). In our zebra finch study, single neurons outperformed behavioral tests in the sense that recognition at the longest propagation distance in the behavioral test was only achieved after repeated training. Narayan et al. (2007) also found several cases in which neural performance was higher than behavioral performance, and they provided two potential explanations for this observation: (1) these results stem from intersubject variability, both experiments having been made on different animals; (2) the information available in the auditory cortex might be reduced in the course of the downstream readout of information, either because of additional source(s) of noise at the readout stage, or because the readout is able to access the best neurons only after repeated exposures via a learning mechanism.

I support the latter explanation, and propose that this neural substrate for individual discrimination of degraded vocalizations that we found in zebra finches may constitute a neural basis for vocal recognition in auditory scenes, that is, a "pool of discriminability" which could then develop in the secondary auditory areas following the birds' experience for the learned recognition of specific individuals in natural settings. In other words, this small population of highly discriminant neurons could take part in a bottom-up process creating a potential for individual discrimination, while top-down processes such as attentional or reinforcement mechanisms would develop this potential through perceptual plasticity. The more numerous, less selective neurons could serve as a pool of cells that can acquire selectivity as the bird learns to discriminate new auditory objects (Meliza et al. 2010). Indeed, these authors reported that during the acquisition of new songs by songbirds in perceptual learning tasks, CMM neurons rapidly changed their selectivity toward the reinforced novel stimuli and then lost this selectivity as the animal learned new stimuli. This "recruitment" of neurons in the wake of experience-dependent needs converges on the concept of learning threshold developed by Seitz and Dinse (2007), in which standard neural responses must be boosted by plasticity-inducing factors such as attention or signal reinforcement, or optimized sensory inputs such as multisensory stimulation, to drive the system past the point of a learning threshold and induce augmented neural responses, and potentially lead to a change in behavior.

8.6 Conclusion and Perspectives

In this chapter, I have endeavored to take a transversal view on the transmission of fine-detailed auditory information at long distance, from sender to receiver, focusing on a songbird model and emphasizing the use of stimuli and paradigms that reflected the behavioral relevance of the challenges that these animals face in their natural environment. While the individual vocal signature encoded in the distance calls of zebra finches is remarkably resistant to propagation-induced sound degradation, a number of neurons in the auditory cortex have the spontaneous ability to encode individual identity of previously unknown vocalizers, as well as propagation distance, in extremely degraded calls. This latter finding is particularly compelling since the subjects had no prior experience in hearing propagated vocalizations. Gentner and Margoliash (2003) suggested that "for vocal recognition, the predictability imparted by species-specific characteristics of vocalizations, and the constraints imposed by evolutionary history and experience, probably yield a population of neurons predisposed to represent those vocalizations." I suspect that the same process, led by evolution, can explain the proficiency with which some neurons deal with propagation-induced degradation.

Regarding the neural encoding of auditory information, the field of neurophysiology is turning toward the investigation of neural networks, following the idea that population coding drives behavior (Curto et al. 2009; Panzeri et al. 2015). One may argue that, seeing the link that was found between single neuron performance and behavior, the known sparseness of neural coding representing the natural world, and the proposed idea of a neural substrate with a potential for high-level discrimination, analyzing single neuron encoding properties is still critical to understanding the link between neural and behavioral levels (Narayan et al. 2006). There is however a general consensus for the idea that although single neurons might drive percepts and elicit behavioral responses, everyday perception is probably the result of coordinated activity by neural populations (Panzeri et al. 2010). Recently, Panzeri et al. (2017) made the case that the two processes underlying perceptual decisions (that is, how neural responses encode stimuli and how they inform behavioral choices) have mainly been studied separately, the latter one having received much less attention. They proposed to redefine the neural code as the neural features that carry sensory information used by the animal to drive appropriate behavior, since only the features that lie in the intersection between sensory coding and information readout can be used to convert sensory perception into appropriate behavioral actions. This interesting approach would help scientists target more efficiently the processes that are actually used by animals in their reactions to their sensory environment. In the case of our focal study, this investigation method, coupled with the computation of the neurons' spectro-temporal receptive fields (STRFs; Theunissen et al. 2001; Nagel et al. 2011; Meyer et al. 2015), could help pinpoint which acoustic parameters in the degraded calls are actually used in the neural code driving behavioral responses. It would be interesting to be able to compare the acoustic parameters that were found to be important for individual discrimination in the acoustic analysis with the acoustic features that birds use while performing this task. This would imply for example the recording of neural responses in a freely moving animal performing an auditory discrimination task in an operant conditioning setup, so as to link auditory input to behavioral output.

In this chapter I have focused on the transmission of individual signatures that were degraded by propagation, combining factors such as sound attenuation, spectro-temporal degradations, and ambient noise in the analysis. In their daily lives however, animals encounter other aspects of auditory scene analysis as well, such as conspecific noise, directionality and auditory source separation (Lewicki et al. 2014). In our case, the use of a loudspeaker to broadcast sounds during electrophysiological recordings effectively annihilated any spatial information that our subjects would have had access to in a natural environment. This probably lead us to underestimate the natural discriminability rather than the opposite, since animals are able to use spatially separated sound sources to enhance their discrimination performance (Dent et al. 2009; Maddox et al. 2012). In any case, combining more of these factors in future research, while investigating the interaction between neurophysiological processes at the brain level and the behavioral output of the animal, will improve our understanding of how animals process real-world scenes such as mate recognition in the natural environment, and bring further insight into how the brain processes this seemingly easy and yet computationally challenging task that is auditory scene analysis.

References

Ahissar M, Hochstein S (1993) Attentional control of early perceptual learning. Proc Natl Acad Sci 90:5718–5722

Appeltants D, Gentner TQ, Hulse SH et al (2005) The effect of auditory distractors on song discrimination in male canaries (*Serinus canaria*). Behav Process 69:331–341. https://doi.org/10.1016/j.beproc.2005.01.010

Aubin T, Jouventin P (2002) How to vocally identify kin in a crowd: the penguin model. In: Advances in the study of behavior. Elsevier, Amsterdam, pp 243–277

Aubin T, Mathevon N, Silva ML et al (2004) How a simple and stereotyped acoustic signal transmits individual information: the song of the white-browed warbler *Basileuterus leucoblepharus*. An Acad Bras Cienc 76:335–344

Bathellier B, Ushakova L, Rumpel S (2012) Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron 76:435–449. https://doi.org/10.1016/j.neuron.2012.07.008

Beckers GJL, Gahr M (2012) Large-scale synchronized activity during vocal deviance detection in the zebra finch auditory forebrain. J Neurosci 32:10594–10608. https://doi.org/10.1523/JNEUROSCI.6045-11.2012

Billimoria CP, Kraus BJ, Narayan R et al (2008) Invariance and sensitivity to intensity in neural discrimination of natural sounds. J Neurosci 28:6304–6308. https://doi.org/10.1523/JNEUROSCI.0961-08.2008

Boumans T, Vignal C, Smolders A et al (2008) Functional magnetic resonance imaging in zebra finch discerns the neural substrate involved in segregation of conspecific song from background noise. J Neurophysiol 99:931–938. https://doi.org/10.1152/jn.00483.2007

Bregman AS (1993) Auditory scene analysis: hearing in complex environments. In: McAdams SE, Bigand EE (eds) Thinking in sound: the cognitive psychology of human audition. Clarendon Press/Oxford University Press, Oxford

Brémond J-C, Aubin T (1990) Responses to distress calls by black-headed gulls, *Larus ridibundus*: the role of non-degraded features. Anim Behav 39:503–511

Brenowitz EA (1982) The active space of red-winged blackbird song. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 147:511–522

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12:4745–4765

Brumm H, Naguib M (2009) Environmental acoustics and the evolution of bird song. In: Naguib M, Janik VM, Zuberbühler K, Clayton NS (eds) Advances in the study of behavior. Academic, New York, pp 1–33

Brumm H, Todt D (2002) Noise-dependent song amplitude regulation in a territorial songbird. Anim Behav 63:891–897. https://doi.org/10.1006/anbe.2001.1968

Capsius B, Leppelsack H-J (1996) Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. Hear Res 96:59–70. https://doi.org/10.1016/0378-5955(96)00038-X

Caras ML, Sanes DH (2017) Top-down modulation of sensory cortex gates perceptual learning. Proc Natl Acad Sci 114:9972–9977. https://doi.org/10.1073/pnas.1712305114

Chew SJ, Vicario DS, Nottebohm F (1996) A large-capacity memory system that recognizes the calls and songs of individual birds. Proc Natl Acad Sci 93:1950–1955

Christison-Lagay KL, Gifford AM, Cohen YE (2015) Neural correlates of auditory scene analysis and perception. Int J Psychophysiol 95:238–245. https://doi.org/10.1016/j.ijpsycho.2014.03.004

Covey E (2000) Neural population coding and auditory temporal pattern analysis. Physiol Behav 69:211–220. https://doi.org/10.1016/S0031-9384(00)00203-1

Cunningham JP, Yu BM (2014) Dimensionality reduction for large-scale neural recordings. Nat Neurosci 17:1500–1509. https://doi.org/10.1038/nn.3776

Curto C, Sakata S, Marguet S et al (2009) A simple model of cortical dynamics explains variability and state dependence of sensory responses in urethane-anesthetized auditory cortex. J Neurosci 29:10600–10612. https://doi.org/10.1523/JNEUROSCI.2053-09.2009

Cynx J, Lewis R, Tavel B, Tse H (1998) Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata*. Anim Behav 56:107–113. https://doi.org/10.1006/anbe.1998.0746

Dahmen JC, King AJ (2007) Learning to hear: plasticity of auditory cortical processing. Curr Opin Neurobiol 17:456–464. https://doi.org/10.1016/j.conb.2007.07.004

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

Dent ML, McClaine EM, Best V et al (2009) Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). J Comp Psychol 123:357–367. https://doi.org/10.1037/a0016898

Einevoll GT, Kayser C, Logothetis NK, Panzeri S (2013) Modelling and analysis of local field potentials for studying the function of cortical circuits. Nat Rev Neurosci 14:770–785. https://doi.org/10.1038/nrn3599

Fay RR, Popper AN (2000) Evolution of hearing in vertebrates: the inner ears and processing. Hear Res 149:1–10. https://doi.org/10.1016/S0378-5955(00)00168-4

Feng AS, Ratnam R (2000) Neural basis of hearing in real-world situations. Annu Rev Psychol 51:699–725. https://doi.org/10.1146/annurev.psych.51.1.699

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hear Res 151:167–187. https://doi.org/10.1016/S0378-5955(00)00224-0

Fitzpatrick DC, Batra R, Stanford TR, Kuwada S (1997) A neuronal population code for sound localization. Nature 388:871–874

Forrest TG (1994) From sender to receiver: propagation and environmental effects on acoustic signals. Am Zool 34:644–654

Fotheringham J (1997) Song transmission and auditory perception of distance in wood warblers (Parulinae). Anim Behav 53:1271–1285. https://doi.org/10.1006/anbe.1996.0361

Gaucher Q, Huetz C, Gourévitch B et al (2013) How do auditory cortex neurons represent communication sounds? Hear Res 305:102–112. https://doi.org/10.1016/j.heares.2013.03.011

Gentner TQ (2004) Neural systems for individual song recognition in adult birds. Ann N Y Acad Sci 1016:282–302. https://doi.org/10.1196/annals.1298.008

Gentner TQ, Margoliash D (2003) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674. https://doi.org/10.1038/nature01731

Gentner TQ, Hulse SH, Ball GF (2004) Functional differences in forebrain auditory regions during learned vocal recognition in songbirds. J Comp Physiol A 190:1001–1010. https://doi.org/10.1007/s00359-004-0556-x

Gilbert CD, Sigman M, Crist RE (2001) The neural basis of perceptual learning. Neuron 31:681–697

Gill P, Woolley SMN, Fremouw T, Theunissen FE (2008) What's that sound? auditory area CLM encodes stimulus surprise, not intensity or intensity changes. J Neurophysiol 99:2809–2820. https://doi.org/10.1152/jn.01270.2007

Gold JI, Watanabe T (2010) Perceptual learning. Curr Biol 20:R46–R48. https://doi.org/10.1016/j.cub.2009.10.066

Grace JA, Amin N, Singh NC, Theunissen FE (2003) Selectivity for conspecific song in the zebra finch auditory forebrain. J Neurophysiol 89:472–487

Guyomarc'h J-C, Aupiais A, Guyomarc'h C (1998) Individual differences in the long-distance vocalizations used during pair bonding in European quail (*Coturnix coturnix*). Ethol Ecol Evol 10:333–346. https://doi.org/10.1080/08927014.1998.9522847

Hauber ME, Cassey P, Woolley SMN, Theunissen FE (2007) Neurophysiological response selectivity for conspecific songs over synthetic sounds in the auditory forebrain of non-singing female songbirds. J Comp Physiol A 193:765–774. https://doi.org/10.1007/s00359-007-0231-0

Holland J (1998) Degradation of wren *Troglodytes troglodytes* song: implications for information transfer and ranging. J Acoust Soc Am 103:2154. https://doi.org/10.1121/1.421361

Hsu A, Woolley SMN, Fremouw TE, Theunissen FE (2004) Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. J Neurosci 24:9201–9211. https://doi.org/10.1523/JNEUROSCI.2449-04.2004

Huetz C, Gourévitch B, Edeline J-M (2011) Neural codes in the thalamocortical auditory system: from artificial stimuli to communication sounds. Hear Res 271:147–158. https://doi.org/10.1016/j.heares.2010.01.010

Jeanne JM, Thompson JV, Sharpee TO, Gentner TQ (2011) Emergence of learned categorical representations within an auditory forebrain circuit. J Neurosci 31:2595–2606. https://doi.org/10.1523/JNEUROSCI.3930-10.2011

Jen PH-S, Chen QC, Sun XD (1998) Corticofugal regulation of auditory sensitivity in the bat inferior colliculus. J Comp Physiol A 183:683–697. https://doi.org/10.1007/s003590050291

Karino G, George I, Loison L et al (2016) Anesthesia and brain sensory processing: impact on neuronal responses in a female songbird. Sci Rep 6:39143. https://doi.org/10.1038/srep39143

Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron 61:597–608. https://doi.org/10.1016/j.neuron.2009.01.008

Klump GM (1996) Bird communication in the noisy world. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Cornell University Press, Ithaca, pp 321–338

Lewicki MS, Olshausen BA, Surlykke A, Moss CF (2014) Scene analysis in the natural environment. Front Psychol 5:199. https://doi.org/10.3389/fpsyg.2014.00199

Lohr B, Wright TF, Dooling RJ (2003) Detection and discrimination of natural calls in masking noise by birds: estimating the active space of a signal. Anim Behav 65:763–777. https://doi.org/10.1006/anbe.2003.2093

Lopes-dos-Santos V, Panzeri S, Kayser C et al (2015) Extracting information in spike time patterns with wavelets and information theory. J Neurophysiol 113:1015–1033. https://doi.org/10.1152/jn.00380.2014

Maddox RK, Billimoria CP, Perrone BP et al (2012) Competing sound sources reveal spatial effects in cortical processing. PLoS Biol 10:e1001319. https://doi.org/10.1371/journal.pbio.1001319

Mahmud M, Vassanelli S (2016) Processing and analysis of multichannel extracellular neuronal signals: state-of-the-art and challenges. Front Neurosci 10:62. https://doi.org/10.3389/fnins.2016.00248

Margoliash D, Fortune ES, Sutter M et al (1994) Distributed representation in the song system of Oscines: evolutionary implications and functional consequences. Brain Behav Evol 44:247–264

Marler P, Slabbekoorn H (2004) Nature's music, the science of birdsong. Elsevier, London

Mathevon N, Aubin T (1997) Reaction to conspecific degraded song by the wren *Troglodytes troglodytes*: territorial response and choice of song post. Behav Process 39:77–84. https://doi.org/10.1016/S0376-6357(96)00046-0

Mathevon N, Aubin T, Brémond J-C (1997) Propagation of bird acoustic signals: comparative study of starling and blackbird distress calls. C R Acad Sci III Sci Vie 320:869–876. https://doi.org/10.1016/S0764-4469(97)80871-2

Mathevon N, Aubin T, Vielliard J et al (2008) Singing in the rain forest: how a tropical bird song transfers information. PLoS One 3:e1580. https://doi.org/10.1371/journal.pone.0001580

Meliza CD, Margoliash D (2012) Emergence of selectivity and tolerance in the avian auditory cortex. J Neurosci 32:15158–15168. https://doi.org/10.1523/JNEUROSCI.0845-12.2012

Meliza CD, Chi Z, Margoliash D (2010) Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. J Neurophysiol 103:1195–1208. https://doi.org/10.1152/jn.00464.2009

Meyer AF, Diepenbrock J-P, Ohl FW, Anemüller J (2015) Fast and robust estimation of spectro-temporal receptive fields using stochastic approximations. J Neurosci Methods 246:119–133. https://doi.org/10.1016/j.jneumeth.2015.02.009

Middlebrooks JC, Simon JZ (2017) Ear and brain mechanisms for parsing the auditory scene. In: Middlebrooks JC, Simon JZ, Popper AN, Fay RR (eds) The auditory system at the cocktail party. Springer, Cham, pp 1–6

Miller LM, Recanzone GH (2009) Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. Proc Natl Acad Sci 106:5931–5935. https://doi.org/10.1073/pnas.0901023106

Mizrahi A, Shalev A, Nelken I (2014) Single neuron and population coding of natural sounds in auditory cortex. Curr Opin Neurobiol 24:103–110. https://doi.org/10.1016/j.conb.2013.09.007

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. https://doi.org/10.1371/journal.pcbi.1002942

Morton ES (1975) Ecological sources of selection on avian sounds. Am Nat 109:17–34

Mouterde SC, Elie JE, Theunissen FE, Mathevon N (2014a) Learning to cope with degraded sounds: female zebra finches can improve their expertise in discriminating between male voices at long distances. J Exp Biol 217:3169–3177. https://doi.org/10.1242/jeb.104463

Mouterde SC, Theunissen FE, Elie JE et al (2014b) Acoustic communication and sound degradation: how do the individual signatures of male and female zebra finch calls transmit over distance? PLoS One 9:e102842. https://doi.org/10.1371/journal.pone.0102842

Mouterde SC, Elie JE, Mathevon N, Theunissen FE (2017) Single neurons in the avian auditory cortex encode individual identity and propagation distance in naturally degraded communication calls. J Neurosci 37:3491–3510. https://doi.org/10.1523/JNEUROSCI.2220-16.2017

Nagel K, Kim G, McLendon H, Doupe A (2011) A bird brain's view of auditory processing and perception. Hear Res 273:123–133. https://doi.org/10.1016/j.heares.2010.08.008

Naguib M (1995) Auditory distance assessment of singing conspecifics in Carolina wrens: the role of reverberation and frequency-dependent attenuation. Anim Behav 50:1297–1307

Naguib M, Hammerschmidt K, Wirth J (2001) Microgeographic variation, habitat effects and individual signature cues in calls of chiffchaffs *Phylloscopus collybita canarensis*. Ethology 107:341–355

Narayan R, Graña G, Sen K (2006) Distinct time scales in cortical discrimination of natural sounds in songbirds. J Neurophysiol 96:252–258. https://doi.org/10.1152/jn.01257.2005

Narayan R, Best V, Ozmeral E et al (2007) Cortical interference effects in the cocktail party problem. Nat Neurosci 10:1601–1607. https://doi.org/10.1038/nn2009

Nieder A, Klump GM (1999) Adjustable frequency selectivity of auditory forebrain neurons recorded in a freely moving songbird via radiotelemetry. Hear Res 127:41–54. https://doi.org/10.1016/S0378-5955(98)00179-8

Olshausen B, Field D (2004) Sparse coding of sensory inputs. Curr Opin Neurobiol 14:481–487. https://doi.org/10.1016/j.conb.2004.07.007

Pachitariu M, Lyamzin DR, Sahani M, Lesica NA (2015) State-dependent population coding in primary auditory cortex. J Neurosci 35:2058–2073. https://doi.org/10.1523/JNEUROSCI.3318-14.2015

Panzeri S, Brunel N, Logothetis NK, Kayser C (2010) Sensory neural codes using multiplexed temporal scales. Trends Neurosci 33:111–120. https://doi.org/10.1016/j.tins.2009.12.001

Panzeri S, Macke JH, Gross J, Kayser C (2015) Neural population coding: combining insights from microscopic and mass signals. Trends Cogn Sci 19:162–172. https://doi.org/10.1016/j.tics.2015.01.002

Panzeri S, Harvey CD, Piasini E et al (2017) Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. Neuron 93:491–507. https://doi.org/10.1016/j.neuron.2016.12.036

Pinaud R, Terleph TA (2008) A songbird forebrain area potentially involved in auditory discrimination and memory formation. J Biosci 33:145–155. https://doi.org/10.1007/s12038-008-0030-y

Pohl NU, Leadbeater E, Slabbekoorn H et al (2012) Great tits in urban noise benefit from high frequencies in song detection and discrimination. Anim Behav 83:711–721. https://doi.org/10.1016/j.anbehav.2011.12.019

Polley DB, Steinberg EE, Merzenich MM (2006) Perceptual learning directs auditory cortical map reorganization through top-down influences. J Neurosci 26:4970–4982. https://doi.org/10.1523/JNEUROSCI.3771-05.2006

Quian Quiroga R, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci 10:173–185. https://doi.org/10.1038/nrn2578

Rabinowitz NC, Willmore BDB, King AJ, Schnupp JWH (2013) Constructing noise-invariant representations of sound in the auditory pathway. PLoS Biol 11:e1001710. https://doi.org/10.1371/journal.pbio.1001710

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13:87–103

Rek P, Osiejuk TS (2011) No male identity information loss during call propagation through dense vegetation: the case of the corncrake. Behav Process 86:323–328

Romo R, Salinas E (2003) Cognitive neuroscience: flutter discrimination: neural codes, perception, memory and decision making. Nat Rev Neurosci 4:203–218. https://doi.org/10.1038/nrn1058

Schmidt MF, Konishi M (1998) Gating of auditory responses in the vocal control system of awake songbirds. Nat Neurosci 1:513–518

Schneider DM, Woolley SMN (2013) Sparse and background-invariant coding of vocalizations in auditory scenes. Neuron 79:141–152. https://doi.org/10.1016/j.neuron.2013.04.038

Schnitzler H-U, Flieger E (1983) Detection of oscillating target movements by echolocation in the Greater Horseshoe bat. J Comp Physiol A 153:385–391. https://doi.org/10.1007/BF00612592

Schumacher JW, Schneider DM, Woolley SMN (2011) Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. J Neurophysiol 106:500–514. https://doi.org/10.1152/jn.01072.2010

Seitz AR, Dinse HR (2007) A common framework for perceptual learning. Curr Opin Neurobiol 17:148–153. https://doi.org/10.1016/j.conb.2007.02.004

Shamma SA, Elhilali M, Micheyl C (2011) Temporal coherence and attention in auditory scene analysis. Trends Neurosci 34:114–123. https://doi.org/10.1016/j.tins.2010.11.002

Slabbekoorn H (2013) Songs of the city: noise-dependent spectral plasticity in the acoustic phenotype of urban birds. Anim Behav 85:1089–1099. https://doi.org/10.1016/j.anbehav.2013.01.021

Stripling R, Milewski L, Kruse AA, Clayton DF (2003) Rapidly learned song-discrimination without behavioral reinforcement in adult male zebra finches (*Taeniopygia guttata*). Neurobiol Learn Mem 79:41–50

Theunissen FE, Elie JE (2014) Neural processing of natural sounds. Nat Rev Neurosci 15:355–366. https://doi.org/10.1038/nrn3731

Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: a rigorous definition. J Comput Neurosci 2:149–162. https://doi.org/10.1007/BF00961885

Theunissen FE, Shaevitz SS (2006) Auditory processing of vocal sounds in birds. Curr Opin Neurobiol 16:400–407. https://doi.org/10.1016/j.conb.2006.07.003

Theunissen FE, David SV, Singh NC et al (2001) Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Netw Comput Neural Syst 12:289–316

Theunissen FE, Amin N, Shaevitz SS et al (2004) Song selectivity in the song system and in the auditory forebrain. Ann N Y Acad Sci 1016:222–245. https://doi.org/10.1196/annals.1298.023

Vicario DS, Naqvi NH, Raksin JN (2001) Sex differences in discrimination of vocal communication signals in a songbird. Anim Behav 61:805–817. https://doi.org/10.1006/anbe.2000.1651

Vignal C, Mathevon N, Mottin S (2004) Audience drives male songbird response to partner's voice. Nature 430:448–451. https://doi.org/10.1038/nature02645

Vignal C, Mathevon N, Mottin S (2008) Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. Behav Process 77:191–198. https://doi.org/10.1016/j.beproc.2007.09.003

von der Emde G (2004) Distance and shape: perception of the 3-dimensional world by weakly electric fish. J Phys Paris 98:67–80. https://doi.org/10.1016/j.jphysparis.2004.03.013

Wang L, Narayan R, Grana G et al (2007) Cortical discrimination of complex natural stimuli: can single neurons match behavior? J Neurosci 27:582–589. https://doi.org/10.1523/JNEUROSCI.3699-06.2007

Wiley RH, Richards DG (1982) Adaptations for acoustic communication in birds: sound transmission and signal detection. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds. Academic, New York, pp 131–181

Woolley SMN (2006) Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. J Neurosci 26:2499–2512. https://doi.org/10.1523/JNEUROSCI.3731-05.2006

Woolley SMN, Gill PR, Fremouw T, Theunissen FE (2009) Functional groups in the avian auditory system. J Neurosci 29:2780–2793. https://doi.org/10.1523/JNEUROSCI.2042-08.2009

Zann R (1984) Structural variation in the zebra finch distance call. Z Tierpsychol 66:328–345. https://doi.org/10.1111/j.1439-0310.1984.tb01372.x

Zann RA (1996) The zebra finch: a synthesis of field and laboratory studies. Oxford University Press, Oxford