Long-range discrimination of individual vocal signatures by a songbird : from propagation constraints to neural substrate

Solveig Mouterde

To cite this version:

Solveig Mouterde. Long-range discrimination of individual vocal signatures by a songbird : from propagation constraints to neural substrate. Animal biology. Université Jean Monnet - Saint-Etienne, 2014. English. <NNT : 2014STET4012>. <tel-01171698>

HAL Id: tel-01171698
https://tel.archives-ouvertes.fr/tel-01171698
Submitted on 6 Jul 2015

THÈSE
présentée devant
L'UNIVERSITÉ JEAN MONNET de SAINT-ÉTIENNE
pour l'obtention du
DIPLÔME DE DOCTORAT

Soutenue publiquement le 24 Juin 2014
par Solveig MOUTERDE

Discrimination à longue distance des signatures vocales individuelles chez un oiseau chanteur :

Des contraintes de propagation au substrat neuronal

Directeurs de thèse :
Nicolas MATHEVON & Frédéric THEUNISSEN

Composition du jury :

Manfred GAHR : Professeur, Max Planck Institute for Ornithology	Rapporteur
Ole LARSEN : Professeur, University of Southern Denmark	Rapporteur
Thierry AUBIN : Directeur de Recherche au CNRS, Université Paris Sud	Examinateur
Isabelle GEORGES : Chargée de Recherche au CNRS, Université Rennes 1	Examinatrice
Clémentine VIGNAL : Professeur, Université de Saint-Etienne, IUF	Examinatrice
Nicolas MATHEVON : Professeur, Université de Saint-Etienne, IUF	Directeur de thèse
Frédéric THEUNISSEN : Professeur, University of California Berkeley	Directeur de thèse

Long-range discrimination of individual vocal signatures by a songbird:
From propagation constraints to neural substrate

Discrimination à longue distance des signatures vocales individuelles chez un oiseau chanteur :
Des contraintes de propagation au substrat neuronal

This PhD was financially supported by a grant from the French Ministry of Research as well as a Monahan fellowship and a Fulbright fellowship to Solveig Mouterde, and fundings from the France-Berkeley Fund, the Agence Nationale de la Recherche (ANR), the University of Saint-Etienne and the National Institute of Health (NIH).

Cover illustration: David Mackie.

Publications

Entre le dépôt de ce manuscrit et la soutenance de thèse, les manuscrits correspondant aux deux premiers chapitres de ce travail ont été acceptés pour la publication.

Voici donc les références bibliographiques les plus à jour au moment de l'impression de cet ouvrage :

Mouterde, S.C. *et al.* (2014). Acoustic communication and sound degradation: How do the individual signatures of male and female zebra finch calls transmit over distance? *PLOS ONE*, in press.

Mouterde, S.C. *et al.* (2014). Learning to cope with degraded sounds: Female zebra finches can improve their expertise at discriminating between male voices at long distance. *The Journal of Experimental Biology*, in press.

Remerciements - Acknowledgements

Ma reconnaissance va en premier lieu à mes deux directeurs de thèse, **Nicolas Mathevon** et **Frédéric Theunissen**, que je remercie de m'avoir fait confiance et de m'avoir permis de poser un deuxième pied dans la recherche. J'ai vraiment eu beaucoup de chance d'avoir comme directeurs de thèse deux personnes dont les qualités humaines égalent la rigueur professionnelle. Merci pour votre bonne humeur, votre disponibilité, et pour nos conversations sur des sujets scientifiques ou plus personnels. Vous êtes pour moi une source d'inspiration sur tous les plans.

Je remercie les membres du jury qui me font l'honneur d'évaluer ma thèse : **Manfred Gahr**, **Ole Larsen**, **Thierry Aubin**, **Isabelle Georges** et **Clémentine Vignal**. J'ai hâte de vous présenter mon travail.

Il est une personne qui a été pour moi une bouée de secours dans des moments difficiles, et d'une aide précieuse tout au long de ma thèse : **Julie Elie**. Ta disponibilité, ton énergie et l'aide que tu donnes sans compter ont été une chance pour moi. Merci pour tout.

Un grand merci à **Colette Bouchut**, pour être celle qui peut trouver une solution à (presque !) tous les problèmes, et ce dans la bonne humeur ! Merci pour ton aide et ta gentillesse, dans et en dehors du labo. J'écris ce texte assise sur une de tes chaises ! Merci à **Nicolas Boyer** pour ton énergie positive et ravageuse, tu es le grand bol d'air frais du labo. Avoir vu la vidéo de toi chantant au volant d'une voiture sur une piste Australienne m'a fait regretter de ne pas être partie sur le terrain !

Je remercie la "team oiseau", **Clémentine Vignal**, pour son sourire et ses conversations passionnées à la table du déjeuner, **Avelyne Villain**, pour être la délicieuse personne que tu es ; je te souhaite tout le meilleur pour ta thèse. Merci à **Emilie Perez**, pour m'avoir montré la "voie du thésard", en passant sa thèse quelques mois avant moi, et pour m'avoir aidée à mettre le pied dans l'équipe oiseau, en me montrant comment baguer mon premier diamant mandarin. Merci à **Ingrid Boucaud** pour sa gentillesse et son calme à toute épreuve (sauf quand il n'y a plus de gâteau). Bon courage pour ta dernière année !

Merci à **Florence Levréro**, pour ton amitié précieuse, ton soutien et ton aide dans les derniers jours de rush ; je sais que l'on restera en contact où que j'atterrisse après la thèse, on n'en a pas fini de s'apprécier ! Merci à **Alain Blanc**, mon "voisin de bureau", pour avoir été le témoin bienveillant de mes 'craquages' de thésarde au bout du rouleau. J'ai été ravie de partager mon amour immodéré du chocolat avec toi. Merci à **Marilyn Beauchaud** pour ta gentillesse et ta bonne humeur à toute épreuve, et **à Joël Attia**, pour les fous rires, les frites partagées et les virées au marché.

Je remercie également les autres membres du laboratoire ENES, passés ou présents, qui ont contribué à faire régner la bonne ambiance au labo : **Frédéric Sèbe**, merci pour nos conversations à bâtons rompus et jusqu'à une heure tardive sur mon avenir, ce que c'est d'être postdoc, ou encore comment préparer sa retraite... Tu es toujours là quand il faut, et j'espère de tout cœur que ton souhait professionnel va se réaliser. Merci à **Mylène Mariette**, **Fabrice Dentressangle** et **Erik Gustafsson** pour votre amitié, votre gentillesse, et nos chouettes soirées chez les uns et les autres. Merci à **Alexis Koutseff** pour son humour hors normes, à **Frédéric Bertucci** et **Delphine Scaion** pour leur gentillesse, et à **Sumir Keenan**, ma voisine de bureau, pour nos conversations qui donnent envie de voyager. Bonne chance pour ta thèse, par monts et par vaux ! Merci également à **Hédi Soula** de m'avoir fait découvrir, indirectement certes, le très utile Screen ! J'ai aussi une pensée pour les stagiaires de plus ou moins longue durée qui sont passés dans ce labo, et qui m'ont vue dernièrement me retrancher de plus en plus dans mon bureau à l'approche de la date de remise de mon manuscrit : merci en particulier à **Pénélope**,

Thibault, **Sébastien**, **Mathieu** et **Mélissa** pour votre bonne humeur, et vous allez voir, je vais recommencer à laisser la porte ouverte !

J'adresse un grand merci à l'équipe enseignement avec laquelle j'ai eu plaisir à travailler : **Odile Liset**, **Marie Montabonnel**, **Pascal Disle**, **Wanda Pazdzior** et **Anne-Laure Verdonck**. Merci pour votre efficacité ! Dans l'équipe administrative, je remercie tout particulièrement **Corinne Marcon** et **Françoise Largeron**, qui se sont toujours montrées très souriantes et sympathiques, et l'ensemble du personnel de la bibliothèque, toujours très aidant. Merci aussi à **Arhama Abdallah**, pour sa présence discrète et très efficace.

Je passe à présent de l'autre côté de l'Atlantique pour remercier mes collègues du Theunissen lab, qui m'ont été d'un soutien sans faille. I want to thank **Tyler Lee** for his friendship, his coolness in every circumstance and his constant help with programming. I will miss our conversations about science and life, and who will I talk to now about my love-hate relationship with Linux? I am also incredibly grateful to **Mike Schachter** for his amazing kindness and help. Thank you for your support in the difficult times, and for offering your help at all times. Thank you **Wendy de Heer** for your kindness and support, and for the infamous and elusive lollipop! I hope you find your dream job in the Bay Area. And same to you, **Michelle Carney**, wherever your hyperactive mind leads you. Thanks for your help down in NAF, and for the great music band that we started with **Chris Warner**! A thousand 'arigato's to **Yuka Minton**, for your constant help inside and outside the lab, and your warm friendship. I was so glad to go watch the last Miyazaki with you, and meet your cat-lady daughter **Asuka**, amongst other things! I am also grateful for meeting many wonderful people in the neuroscience labs: **Taffeta Elliott**, **Alex Huth**, **Mark Lescroart**, I can't cite you all but it's been great hanging out with you.

Je remercie du fond du cœur ma famille, et notamment **mes parents**, pour leur soutien inconditionnel, et leurs opérations "tornade blanche" dans mon appartement ! Merci à **mon frère** et à **ma sœur** d'être là malgré la distance, et j'ai hâte d'accueillir la surprise du mois d'août !

Je veux également remercier mes amis, qui, de près ou de loin, ont un lien avec cette thèse: **Béatrice**, pour ton amitié et ton soutien sans faille ; **Lise,** pour être toujours là pour moi. **Magali**, pour tes encouragements et tes "Forza !!". **Daphné**, **Ariane**, **Minh**, **Lucie**, **Mushi**, **Eve**, **LN**, **Audrey** et tous mes amis vétos, pour vos amitiés précieuses. **Melinda**, **Mica et Nolwenn**, **Stéphane et Doris**, qui savent ce que c'est que de préparer une thèse ; en particulier, merci mille fois Melinda pour ta présence lors des derniers mois. Merci à celles et ceux qui ont embelli ma vie à Saint-Etienne : **Emilie**, **Mélanie**, **Bérengère**, **Laurent**, **Bertrand**, les deux **David D.**, **Anne-Laure**, la **famille Rosemont**, **Marie-Eve**, **Céline R.** et tous les improvisateurs stéphanois !

Sur la côte Pacifique, je remercie ma chère **Lisa** pour tout, quel bonheur de t'avoir rencontrée ! **Noopur**, you are the most amazing and giving person in the world, I'm so glad to call you my friend. Thank you **Mayur** for your friendship, you support and your lolcats! Merci à **Ela, Clara, Brian**, **Michelle**, **Dev**, **Beatriz**, **Alex**, **Vincent**, **Marion**, **Ségolène** et **Thomas**, **Ihab**, **Michaël**, la **famille Theunissen**, je ne peux pas citer tout le monde mais vous avez fait de mon séjour aux USA une expérience formidable.

Enfin, je veux remercier celui qui partage ma vie, même si la distance n'est pas toujours facile. **David**, thank you from the bottom of my heart for your love, attention and support, day or night, rain or shine. I am thankful for this thesis because this is how I found you.

Et pour finir, je suis bien entendu redevable aux organismes ont accepté de financer cette thèse et de participer à son bon déroulement : le Ministère de l'Enseignement Supérieur et de la Recherche, l'Université Jean Monnet, le Fonds France-Berkeley, l'Agence Nationale de la Recherche (ANR), le National Institute of Health (NIH), et surtout la Fondation Monahan et la Commission Franco-Américaine Fulbright, qui m'ont permise de vivre une expérience inoubliable aux Etats-Unis.

TABLE OF CONTENTS

INTRODUCTION

I. INTRODUCTION

A. Scene analysis: a complex task, surprisingly easy to perform, yet hard to study

Scene analysis, defined as the ability to extract relevant information from a complex environment, is a universal problem solved by all animals in their natural settings (Lewicki et al., 2014). As easy as it may seem to us to discriminate an object from the background or from the surface it stands on, or to discern the voice of someone talking to us in the midst of the many sounds that arrive simultaneously at our ears, and as proficiently as many animals have been shown to perform these same tasks (Appeltants et al., 2005; Aubin and Jouventin, 2002a; Hill, 1979; Schnitzler and Flieger, 1983; von der Emde, 2004), the underlying biological processes involved in scene analysis are very complex (Shamma et al., 2011) and their study has been a major challenge for scientists, especially on the neurophysiological and computational levels.

How the brain that continuously receives tremendous amounts of information from different sources and different perceptual channels manages to discriminate what is relevant from what is not, how it interprets this sensory information and makes sense of the external environment, and how this process results in behavioral responses that are adapted to the situation and context, are biological questions that represent an exciting challenge and many potential insights for scientists studying integrated processes, from behavior to brain.

The term 'scene analysis' was actually first used by researchers in computer vision and later extended to the study of visual perception in animals, investigating how the processing of surfaces, edges, colors or textures could lead to the representation of visual objects (Allen and Higham, 2013; Bregman, 1993). The notion of auditory scene analysis has later followed by analogy, and is defined as the process enabling animals (or computers) to segregate and categorize sounds amidst the multiple sources of noise in natural environments, in order to perceptually organize information about biologically relevant auditory objects (Moss and Surlykke, 2001); the processing of auditory scenes has been suggested to be a driving force in the evolution of vertebrate hearing (Fay and Popper, 2000). Studies in this field have resulted in the development of interesting conceptual and methodological approaches; for example, one prominent hypothesis is that auditory objects segregate into separate "streams" whenever they activate well separated populations of neurons that are selective to particular attributes of these objects (Bregman and Campbell, 1971); this has been investigated with behavioral (MacDougall-Shackleton et al., 1998), computational (McCabe and Denham, 1996) and neurophysiological approaches (Bee and Klump, 2004). Other studies highlighted the fact that auditory scene analysis results from a combination of the neural encoding of acoustic properties of sounds in the ascending auditory pathway ("bottom-up" processes) and higher-order influences such as selective attention, memory or context ("top-down" processes), and emphasized the importance of assessing the extent to which both effects play a role in the neural representation of auditory objects (Alain et al., 2001; Shamma and Micheyl, 2010; Snyder and Alain, 2007; Zekveld et al., 2006).

However, a major limitation in current neurophysiological approaches is that they do not acknowledge or address the complexity of the problems that need to be solved (Lewicki et al., 2014); in order to provide a thorough understanding of the mechanisms in play in complex natural settings, Lewicki *et al.* concluded that research should focus more on the real-life issues faced by animals in their natural environment, using stimuli and paradigms that reflect as much as possible the ecologically relevant stimuli and behaviors of the tested species. The work presented here has been conducted with that challenge in mind, using natural stimuli obtained from real-life environments. Given that birds are "acoustic specialists", and that they have been used as models of investigation in bioacoustics research for decades, I chose to focus on auditory scene analysis in a bird species facing the constraint of communicating at long range. More specifically, I studied how information contained in the vocalizations is degraded through communication at long distance (**manuscript 1**), and how birds extract the information initially encoded in the signals in their degraded state (**manuscripts 2 & 3**).

In the next sections of this introduction I will first describe the process of vocal communication and the constraints imposed on the transmission chain by long distance propagation in natural environments. Next, I will focus on my model species, the zebra finch (*Taeniopygia guttata*), and explain why it was an ideal subject for investigating these questions. Finally, I will detail my research objectives and explain how I conducted this research, navigating between the Universities of Saint-Etienne (France) and Berkeley (USA).

B. The problem of communicating at a distance

1. Vocal communication: transmitting acoustic information

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. Formalizing such a broad concept is not an easy task, and has been the subject of scientific debate; here I define communication by the transmission of information by a sender (or emitter) to a receiver that will use this information to make a decision about how it should respond (Bradbury and Vehrencamp, 1998). The vehicle bearing the information is called the signal, and it may inform the receiver about the sender's specific and individual identity ("individual signature"), sex, social group, age class or physical qualities; it can also inform about the sender's location, its motivation (such as in aggressive, territorial or courtship signals) and its emotional state like stress (Doucet, 2003; Duvall, 1979; Hauser, 1996; Kondo and Watanabe, 2009; Mathevon et al., 2010; Naguib, 1995; Perez et al., 2012).

The communication transmission chain is organized as follows. The emitter encodes information destined for the receiver(s) (*Fig. 1*). The signal then propagates through the transmission canal (air, water, soil...) and its information content is altered as a result of environmental constraints (Shannon and Weaver, 1949). The receiver detecting this modified signal extracts the information it contains, depending on its physiological abilities and the signal's degradation level, and responds accordingly to it. In the case of vocal communication the signal is a sound, which is an acoustic wave, i.e. a disturbance generated in an elastic medium such as air that propagates away from the source. Sounds can be characterized by their amplitude, usually measured in intensity (W/m^2) or pressure (Pa), the frequency content of the wave (Hz) and the temporal dynamics of both parameters (Forrest, 1994).

Thus, two major constraints are at play in the communication process: the physiological abilities of both sender and receiver that regulate respectively the encoding and decoding of information, and the environmental constraints that systematically decrease the information content of the signal. For a given vocalization in a given species, these constraints determine the distance from the source (emitter) over which the signal can be perceived by potential receivers;

this has been termed the active space of a signal (Brenowitz, 1982). It is of fundamental importance in acoustic communication as it pertains to the biological relevance of propagated signals (Marler and Slabbekoorn, 2004). In the next paragraphs I will describe in greater detail the effects of propagation on acoustic signals, and the physiological and behavioral adaptations displayed by the emitters and receivers to counteract these deleterious effects.

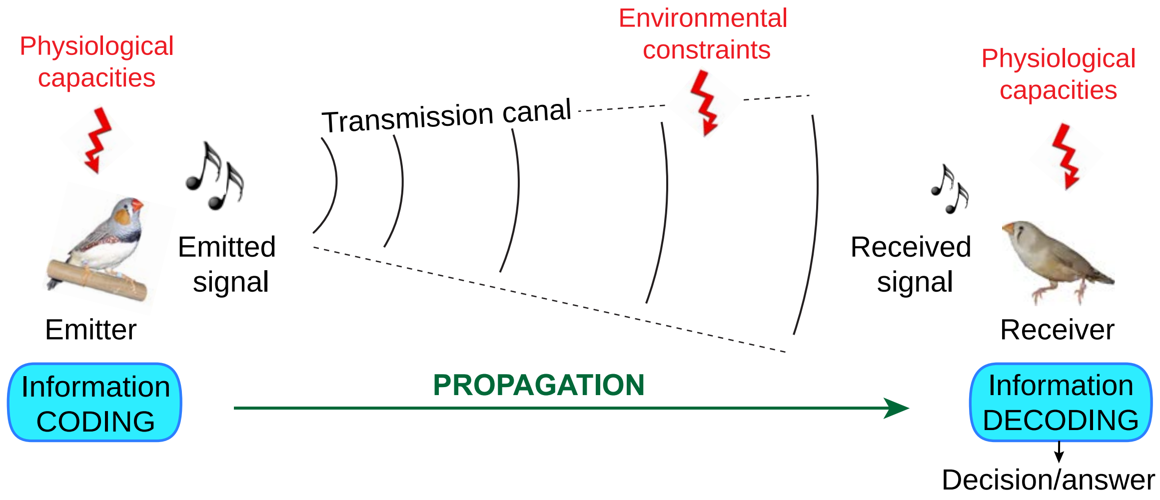


Fig. 1: The chain of information transmission. A signal carrying information is transmitted from an emitter to a receiver. The emitter encodes information in the signal, depending on its physiological capacities. The signal transmits through the environment, and is systematically degraded due to environmental constraints: the signal that is perceived by the receiver contains less information than the initial signal. Depending on its physiological capacities, the receiver decodes this degraded message and reacts accordingly.

2. Sound propagation in the environment and signal degradation

Acoustic signals transmitted over large distances through the atmosphere degrade in amplitude, spectral and temporal structure as they propagate through the environment (Forrest, 1994; Wiley and Richards, 1982). As sound waves radiate from the source, their energy will decrease with distance as a result of geometric spreading. In a completely homogeneous environment, this spherical loss accounts for a standard attenuation of sound pressure of 6 dB per doubling distance around the point source of the sound. Natural environments are not homogeneous however, which causes further degradation in the form of excess attenuation. This additional degradation can be caused by atmospheric absorption (that depends on temperature and humidity), by the scattering of sound waves as they propagate through heterogeneous media (reflection, diffraction and refraction of sound due to physical objects along the path or heterogeneities in the atmosphere such as turbulences) and by boundary interferences (e.g., interference between direct and reflected waves when communicating near the ground). The magnitude of these alterations is a function of the wavelength of the sound with higher frequencies suffering higher degradations (Forrest, 1994; Wiley and Richards, 1982).

In the context of biological communication systems, researchers have investigated if habitats are involved differently in propagation-induced sound degradation. A number of sound transmission experiments have confirmed this with the broadcast of pure tones and bird vocalizations through different environments (Cosens and Falls, 1984; Dabelsteen et al., 1993; Marten and Marler, 1977; Morton, 1975). Some generalities were drawn: scattering from vegetation causes reverberations in closed habitats (e.g., forest) and scattering from atmospheric turbulences is more important in open habitats (e.g., grassland). Corollary of this observation, transmission conditions are more stable in closed habitats than in open ones (Morton, 1975). While in all habitats lower frequencies (1-3 kHz) attenuate less than higher frequencies, open habitats are more subject to the attenuation of low frequencies, and the ground effect is more important (Cosens and Falls, 1984).

These experiments opened a new field of studies investigating to what extent the vocalizations from animals are adapted to the environment they live in. Morton (1975) formulated this as the "acoustic adaptation hypothesis", suggesting that the design of acoustic signals has evolved to maximize their transmission distance. Following his work, a number of studies have shown support of this idea that habitat acoustics play a role in shaping the natural acoustic signals (Blumstein and Turner, 2005; Boncoraglio and Saino, 2007; Brown and Handford, 2000; Ey and Fischer, 2009; Gish and Morton, 2010; Morton, 1975). The little greenbul (*Andropadus virens*), an African bird living in different types of habitats, is a good example of intra-species adaptation to ecological constraints: several features such as maximum and minimum frequency and song note rate showed habitat-dependent variation (Slabbekoorn, 2004). Studying the acoustic characteristics of different habitats enabled inferences on the general acoustic structure of "ideal signals" depending on the habitat type. For example, for best resistance to degradation in an open environment, an ideal call should be short with a large bandwidth and have high frequency modulation rate (Ey and Fischer, 2009). Another preponderant parameter influencing the transmission of acoustic signals through the environment is the height of the sound source: the higher the source, the better the transmission (Dabelsteen et al., 1993; Mathevon et al., 2005). In temperate habitats, emitting height and sound frequency have been found to be more important than the habitat in determining how far the sound will carry (Marten and Marler, 1977).

Finally, signal transmission may also be impaired by ambient environmental noise, which includes sounds generated by physical processes, such as wind or rain, and sound generated by biological organisms, such as other animals' acoustic signals. The level and the spectral and temporal characteristics of background noise are major determinants of the active space of a vocal signal (Klump, 1996). This masking effect can be measured by the signal-to-noise ratio (SNR), which is the relative energy of the signal to the background noise. To this term defined here in the strict sense, I propose and will use a broader level of meaning for the SNR, which takes into account not only the signal's amplitude relative to that of the background noise, but also the degradation of the sound's temporal and spectral structure due to propagation through the environment, as detailed above. In this inclusive definition, the SNR thus describes how much information is left in the signal at a given point.

3. How emitters and receivers adapt to adverse communication conditions

a. The emitter's solutions: How to reach the wanted audience?

Physiological and behavioral adaptations of the emitters to adverse communication conditions have been the subject of numerous studies. Here I give a brief review of the different strategies developed at this end of the transmission chain.

Information coding. Studies on specific and individual identity encoded in vocalizations suggest that not only the acoustic features in the signal are adapted to transmission in the animal's environment, but that the actual coding of information in the signal is generally based on sound features resistant to propagation (Mathevon and Aubin, 1997), and in any case adapted to the biological requirements of the species in its environment. For example, in the white-browed warbler *Basileuterus leucoblepharus*, a territorial species living in the dense environment of the Brazilian tropical forest, species-specific information in the male song is encoded in a resistant acoustic feature that propagates at long distance, further than 100 m. Conversely, the individual signature degrades rapidly during propagation, restricting individual recognition to neighboring territorial males (Mathevon et al., 2008). For this territorial bird, knowing the identity of the neighboring birds is of critical importance in mediating the male's responses to a song emitted in the vicinity: the bird will not respond to a known song coming from the territory of a regular neighbor, but will fly towards the sound source if the song is perceived as being broadcasted by a stranger. Outside the male's territory, 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. Thus, this communication system is well matched to the acoustic constraints of the rainforest and to the ecological requirements of the species.

Penguins are another example that remarkably illustrates the adaptation of information coding in the acoustic signals to the typical requirements of the species. The various species of the Spheniscidae span all levels of territoriality, from the completely non territorial emperor penguin *Aptenodytes forsteri* that walks around during brooding to the burrow-dwelling and fiercely territorial little blue penguin *Eudyptula minor*. As such, finding its mate during the breeding season is much easier for little blue penguins, that can refer to the topographical landmark of their burrow (Mouterde et al., 2012), than for the emperor penguin that has to recognize the voice of its partner among loud clatter occurring in the dense crowd of conspecifics (Aubin et al., 2000). Acoustic analyses comparing the intra versus inter-individual variability in calls in both species have shown that the vocal signature in the emperor penguins is 3 to 4 times more individualized than the signature of the little blue penguin. The same analysis over other penguin species have shown the existence of a linear relationship between the degree of individuality in the calls and the territoriality of the species (Jouventin, 1982), which accounts for the importance for penguins to rely on acoustic cues to recognize each other (Aubin et al., 2000).

Behavioral adaptations. Animals have been shown to adjust their vocalizations to compensate for changes in the background noise. For example, common marmosets *Callithrix jacchus* increase amplitude levels of their vocalizations in response to increased levels of noise (Brumm et al., 2004). This vocal adjustment, called the Lombard effect, has been shown in numerous species including the zebra finch *Taeniopygia guttata* (Cynx et al., 1998). Other examples for vocal adjustments, in the face of anthropogenic noise, include enhancing the frequencies that are not affected by the low-frequency noise (Wood et al., 2006) or increasing the duration of the vocalizations (Foote et al., 2004) This latter adaptation increases the temporal redundancy in the coding of information, and it has been experimentally demonstrated in king penguins (*Aptenodytes patagonicus*) that redundancy increases the probability of signal detection in a high background noise (Aubin and Jouventin, 2002b). As we have seen previously, sounds transmit better when emitted at a height. Animals have also been found to improve the transmission of their signals in this manner, such as in the blue-black grassquit *Volatinia jacarina* that vocalizes while leaping above the top of dense grass (Wilczynski et al., 1989) or in the wren *Troglodytes troglodytes* that flies to a high song post to respond to the playback of an unfamiliar song (Mathevon and Aubin, 1997). Animals also adapt to the temporal variations of background noise: bladder grasshoppers *Bullacris membracioides* emit stridulatory signals at night, when the transmission conditions are ideal (Van Staaden and Römer, 1997), and birds living in urban areas where ambient noise is high during the day have developed the same strategy (Fuller et al., 2007). Finally, research on the emitters' adaptations to atmospheric conditions has led to contradictory results: studies have suggested that birds and mammals could adapt to cyclical atmospheric conditions in their habitats by vocalizing preferably at dawn and evenings (Brown and Handford, 2002; Larom et al., 1997), but this hypothesis has been challenged (Dabelsteen and Mathevon, 2002); in any case, bat species have been shown to change their echolocation call structure across seasons, using longer duration, lower frequency calls in the higher absorption rainy season (Snell-Rood, 2012).

b. The receiver's solutions: How to optimize the gained information?

While a few studies highlighted behavioral adaptations of receivers that increase the signal reception, physiological adaptations of receivers to the alteration of signals have been well studied in the auditory system; however, less is known about their actual information decoding process.

Behavioral adaptations. While we have seen above that the emitter's position played a role in how well the information was transmitted, receivers can also enhance signal reception by strategically choosing their listening post. For instance, the wren perches at a height in response to playback of propagation-degraded song (Holland et al., 2001; Mathevon and Aubin, 1997). Sound transmission experiments using vocalizations of blackcap birds (*Sylvia atriacapilla*) in forest habitat showed that the height of the receiver was more important than the height of the emitter in decreasing sound degradation (Mathevon et al., 2005): choosing high perches in a forest would improve communication on both sides of the transmission chain, with a greater effect on the receiver's side. It was even suggested that the main reason for choosing high singing posts in male blackbirds (*Turdus merula*) is to improve the singer's ability to hear responses to its song (Dabelsteen et al., 1993). King penguins have also been found to adjust their posture in a way that enhances the reliability of signal reception in the noisy colony (Lengagne et al., 1999).

Adaptations in the auditory system to background noise. All vertebrates have been shown to use peripheral frequency analysis (i.e., selective frequency tuning) to break down complex sounds into their frequency components (Fay and Popper, 2000). This frequency filtering system 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 channels. This is especially true for abiotic noise (such as wind) dominated by low frequencies that are often out of the frequency range of vocalizations (Fay and Popper, 2000; Klump, 1996). The minimum SNR that is necessary for the detection of a sound in background noise, called the critical masking ratio, was found to be the smallest in the spectral region which is biologically relevant for the species (Dooling, 1986; Lohr et al., 2003; Okanoya and Dooling, 1987). This intrinsic auditory filtering system may be further adjusted to the characteristics of background noise: in birds, neurophysiological recordings showed that neuronal filters in the auditory system could be sharpened to adapt to the spectrum of background noise (Nieder and Klump, 1999). Additionally, comparisons between auditory thresholds of urban and woodland bird populations showed differences depending on the frequency range of the tested signals, demonstrating the adaptation of the urban population to the particular noise conditions in an urban environment (Pohl et al., 2012). Other studies have linked the masked auditory detection threshold of the receivers to the environmental conditions (reviewed in Patricelli and Blickley, 2006).

Varying the sensitivity of the auditory system to the spectral content of the signal and background noise is not the only tool used by animals to enhance signal extraction: in ambient noise as well as in signals, energy across frequency regions is often coherently modulated in time, and animals have been found to exploit these non-random amplitude fluctuations to extract signals from the background noise. This process, called co-modulation masking release, has been shown in various species of birds and mammals (Branstetter and Finneran, 2008; Klump and Langemann, 1995; Pressnitzer et al., 2001; Verhey et al., 2003).

These various adaptations of the auditory system to the masking effect of noise can lead to high detection performances: for example, king penguin chicks are able to recognize a parental call even if its level is 6 dB below the level of background noise (Aubin and Jouventin, 1998).

Information decoding. Receivers show great flexibility in decoding degraded information: for example, budgerigars (*Melopsittacus undulatus*) tolerate large spectral and temporal alterations in their discrimination of calls (Park and Dooling, 1986). This tolerance to signal alteration in vocal recognition has been shown in various species of mammals and birds (Charrier et al., 2003; Charrier et al., 2009; Vignal et al., 2008); it may be due to a high level of redundancy in the information coding at the level of the emitter, but also to an adapted decoding strategy on the receiver's side. Indeed, coping with degraded signals may lead to specific adaptations in which the scope of perception is broader than the acoustic space of the emitted signals (as defined by the acoustic features of these signals): the received signals being more variable than the emitted signals after transmission in the environment, receivers would gain to not narrowly focus on the acoustic features of the clean signals (Brumm and Naguib, 2009). Such perceptual strategies were described in suboscine birds of the Amazonian forest (Luther and Wiley, 2009). Other strategies involve relying on acoustic features that are less likely to be degraded through propagation. For example, in the black-headed gull (*Lurus ridibundus*), amplitude modulation is not taken into consideration in the decoding process of distress calls; this appears well adapted to transmission at long-range since amplitude modulation is always subject to strong degradation (Brémond and Aubin, 1992). However, the study of perceptual plasticity at the receiver's level has been generally overlooked, 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). On a related subject, one should note that signal degradation is not only an obstacle to animal communication, as receivers can use this degradation to locate the emitter and estimate the propagation distance (a process called "ranging"; Fotheringham, 1997; Holland, 1998; Naguib, 1995).

The study of the information decoding process at the neural level has been the subject of more recent research. Neurophysiological studies in birds have described populations of neurons in the auditory system that could discriminate between songs and which responses were invariant to noise (Moore et al., 2013; Schneider and Woolley, 2013) or to intensity level (Billimoria et al., 2008). Such neurons illustrate adaptations to certain aspects of auditory scene analysis at the neural level.

There is still a lot to discover about how animals process degraded information. More specifically, it is still unknown how songbirds deal with the impact of sound degradation due to long-distance propagation when performing individual vocal discrimination, at the behavioral and the neural levels.

C. Our model: the zebra finch *Taeniopygia guttata*

1. Ecology and biology

Zebra finches (*Fig. 2*) are small gregarious songbirds from subarid regions of Australia. They live in large flocks in open grassy country with a scattering of trees and bushes and feed almost exclusively on grass seeds (Zann, 1996). Partners form strong pair bonds for life and both parents take care of their offspring. Because these birds are opportunistic breeders living in a very unpredictable environment (Zann, 1996), maintaining strong pair bonds between breeding events while living in large fission-fusion groups is of utmost importance. With the flock being constantly on the move for foraging, topographic landmarks may be scarce and partners do not have a fixed nest site to meet each other: using a solid vocal recognition system could avoid partners the cost of losing each other.

Fig. 2: Zebra finches in an aviary. Three males are on the left of the picture, recognizable by their orange cheek patches; a female sits on the right.

Zebra finches are therefore a species of choice to study vocal recognition between mates, from the analysis of their individual vocal signature to the neural substrate of vocal discrimination and recognition in the auditory cortex. They have been widely used in research for various biological questions such as the sensorimotor process in song learning (Bolhuis and

Gahr, 2006; Eales, 1985; Podos et al., 2009), sexual selection (Burley et al., 1996; Simons and Verhulst, 2011), acoustic communication (Ritschard and Brumm, 2012; Vignal et al., 2004a) or neural representation of sounds (Boumans et al., 2008; Menardy et al., 2012; Woolley et al., 2005). I will detail below how these last two areas of research have been a very useful basis for the work presented in this thesis.

2. Vocal repertoire and individual signature

Zebra finches are highly vocal and use different types of vocalizations to communicate; the best known and most studied vocalization is the song that had originally sparked interest in songbirds as a model system for studying vocal learning, since learning to produce song shows striking similarities to learning to produce speech in humans (Doupe and Kuhl, 1999). In the vast majority of songbirds from temperate regions only the males sing (Langmore, 1998), and they learn by listening to their father or a tutor. In zebra finches, males have a unique and stereotyped song composed of various syllables, which is used in the courtship context (Zann, 1996).

Among the other types of zebra finches' vocalizations, one of the most frequently heard is the distance call; this loud call is emitted by both sexes and is used in many situations where birds establish acoustic contact at a distance (Zann, 1996); as described by Richard Zann: "the distance call 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" (Zann, 1984). The distance call is thus the perfect candidate to study the impact of the propagation-induced degradation of a vocalization on its information content, and investigate how the birds of both sexes manage to perform auditory scene analysis in difficult conditions.

The distance call of zebra finches is a complex sound, consisting of a harmonic series (i.e., harmonic bands that are multiples of a fundamental frequency) modulated in frequency as well as amplitude (*Fig. 3*). Distance calls differ between males and females (Vicario *et al.*, 2001; Zann, 1984), the males' fundamental frequency being higher than the females' (typically 650-1000 Hz versus 500-600 Hz) as well as usually being shorter and more frequency-modulated. This difference between sexes is also seen on the developmental stage: males learn parts of their call while the females do not (Forstmeier *et al.*, 2009). It has been shown that the distance call bears an individual signature, and that zebra finches are capable of call-based individual recognition (Vignal et al., 2004a; Vignal et al., 2008; Zann, 1984). Observations of wild zebra finches in their natural habitat support the hypothesis that distance calls allow long distance communication between individuals: the "active space" of these calls has been estimated based on naturalistic observations to be up to 100 meters by Zann (1996); similar conclusions were reached theoretically, using discrimination thresholds for masked signals in this species (Lohr et al., 2003). However, no experimental work has yet investigated the long-range transmission of distance call based individual signature.

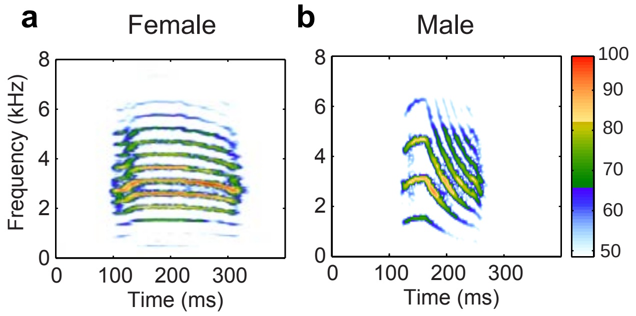


Fig. 3: Spectrograms of a female (a) and male (b) zebra finch distance call. A spectrogram represents the amplitude of the sound as a function of frequency and time (see methods). The color scale is in relative dB as shown on the color bar with 100 dB corresponding to maximum amplitude observed.

Zebra finches thus offer an ideal opportunity for studying how male and female vocal signature is degraded when transmitting through the environment, and how these birds extract this information from degraded signals at the behavioral and neural levels, while working within realistic biological constraints faced by this species.

3. Processing of sounds and auditory scenes in the auditory cortex of zebra finches

In songbirds, auditory information from the sensory system ascends the auditory pathway through the midbrain, thalamus and into the auditory cortex, itself divided into a primary (field L complex) and a secondary region (caudal mesopallium or CM; caudomedial nidopallium or NCM) (*Fig. 4*). Along this ascending pathway, increasing selectivity for natural sounds has been shown (Theunissen et al., 2004). In particular, high selectivity for conspecific songs has been found in the field L, a region that is analogous to the primary auditory cortex of mammals (Grace et al., 2003; Hauber et al., 2007; Theunissen and Shaevitz, 2006), as well as in the CM (Grace et al., 2003; Hsu et al., 2004) and the NCM (Bailey et al., 2002; Ribeiro et al., 1998). Further studies suggested that hierarchical sensory processing occurs in the auditory cortex, with the secondary auditory areas (CM and NCM) showing sensitivity to higher-order features such as behavioral significance of sounds, and playing a role in learned auditory discriminations (Chew et al., 1996; Pinaud and Terleph, 2008). Indeed, neural substrates for individual recognition in the auditory cortex were found in the medial CM (CMM; Gentner and Margoliash, 2003) and the NCM (Chew et al., 1996; Gentner, 2004).

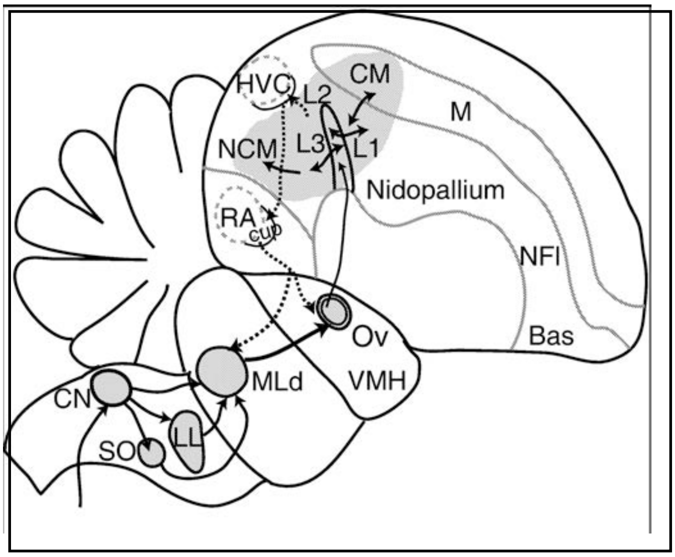


Fig. 4: Auditory pathway in the songbird brain. Auditory sensory input from the hearing organ is relayed to the hindbrain (CN, cochlear nucleus; LL, lateral lemniscus; SO, superior olivary nucleus), followed by the midbrain (MLd, midbrain auditory nucleus) and the thalamus (Ov, ovoidalis), where it projects to the primary auditory cortex, in field L2. The pathway continues to secondary field L zones (L1, L3), then to the secondary auditory areas (CM, caudal mesopallium; NCM, caudal medial nidopallium). HVC (high vocal center) and RA (robust nucleus of the arcopallium) are part of the vocal pathway.

Concerning auditory scene analysis per se, most published research to date has dealt with masked signals and noise or intensity levels, and studies in mammals (Rabinowitz et al., 2013) and songbirds (Vignal et al., 2004b) suggest an increased tolerance for noise (i.e., an increased adaptation to sound statistics resulting in better discrimination of natural sounds embedded in noise) as one ascends the auditory pathway. In songbirds, fMRI-based evidence suggested that the first area within the auditory system where the ability to discern a song from masking noise emerges is located in the NCM (Boumans et al., 2008). This finding was corroborated by two electrophysiology-based studies on zebra finches, the first one studying noise-invariant responses in the NCM using unfamiliar conspecific songs (Moore et al., 2013), and the second one describing a population of neurons in the same region that encoded previously learned individual songs embedded in background chorus noise (recordings in field L did not provide evidence for the existence of such neurons in this area; Schneider and Woolley, 2013). Finally, a study describing a population of field L neurons able to discriminate between bird songs while being invariant to intensity showed that the neurons with a high degree of invariance also displayed a high discrimination performance (Billimoria et al., 2008). This ability would be an important factor of success in the task of discriminating between individuals at a distance in a natural environment.

While recognition of known individual songs and calls has been investigated in the songbird brain, as described above, studies have not directly examined the ability of neurons to discriminate individual identity in unfamiliar vocalizations, in other words, investigated individual vocal discrimination, rather than recognition, at the neural level. Moreover, prior studies involving discrimination of vocalizations used single exemplars of songs, and thus did not directly assess the discrimination of the individual signature in a group of communication calls. The use of calls rather than songs is also a far better choice for this study on propagation-induced degradation since distance calls are used for long distance communication while the directed song is only used in proximity to a female in zebra finches. Besides, compared to calls, songs contain an added layer of information, the syntax (i.e., the arrangement of syllables typical of each male or male lineage), which can also be used in songbirds to perform individual discrimination (Gentner, 2004). Disentangling to which extent the neural processing uses the fine spectro-temporal features of the signal or its syntactic organization to characterize an individual's signature may therefore be challenging when using songs as stimuli. In this light, calls are good candidates for studying vocal discrimination between individuals, especially in difficult transmission conditions; they offer the additional advantage of being able to test vocalizations from both sexes. Finally, the study of the neural representation of auditory scenes, using stimuli combining all aspects of sound degradation in a natural environment, has yet to be published.

D. Research objectives

In this work, I have endeavored to bring my contribution to the study of how animals manage to communicate in complex auditory scenes. More specifically, I have studied how information encoded in vocalizations is altered through long distance propagation, and how the receivers extract and process this degraded information. I chose to focus on the individual identity encoded in vocal signals (the individual signature) and how the fine spectral and temporal acoustical structure needed for this task (Charrier et al., 2003; Latinus et al., 2013; Mathevon et al., 2008) is encoded in the avian auditory system.

Using an animal model that fits the experimental requirements and most importantly has to deal with this issue in its real-life environment, I investigated this question on three levels: first, at the acoustical level, using an analytical approach; second, at the level of the organism, conducting behavioral tests; and third, at the neural level, performing electrophysiological experiments. With this transversal approach, I have ventured to study all three aspects of the transmission chain, from the sender's encoded message to the receiver's processing of the information, via the effects of environmental constraints on the signal's transmission (*Fig. 5*). Throughout the whole process, I used natural calls that had propagated in natural environments and, thus, where the reduction of the SNR was due to the combined effects of intensity decrease, sound degradation and ambient noise. Below I give a brief outline of the methods used in each section of my research study.

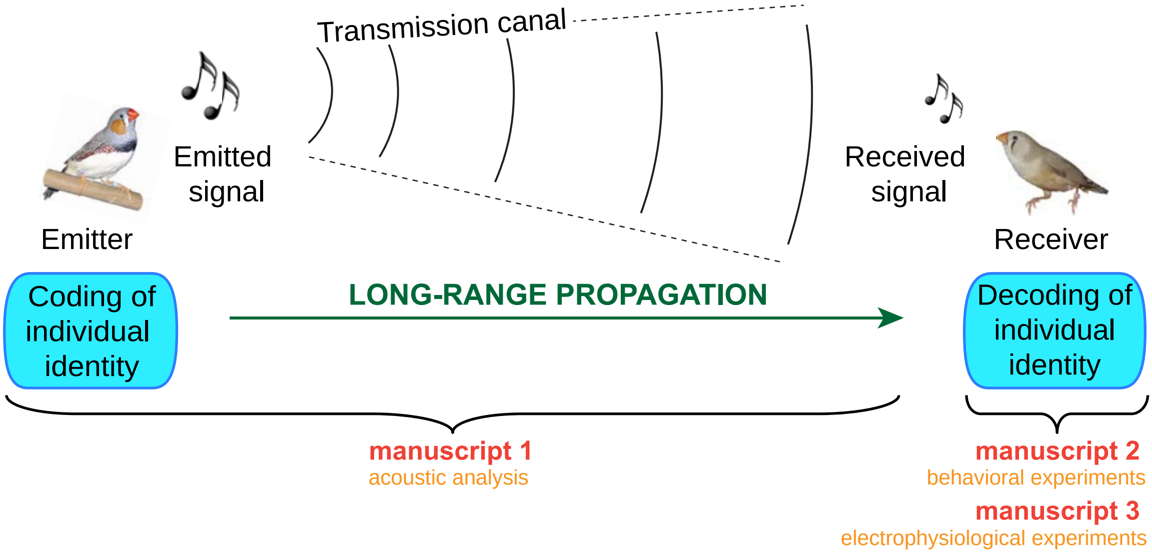


Fig. 5: Research objectives, and how they relate to the information transmission chain. The work presented in manuscript 1 dealt with the coding of information about individual identity and how this information was degraded through propagation. The studies presented in manuscript 2 and 3 dealt with the information decoding by the receiver, either at the behavioral or the neural level.

1. Acoustic analysis of the individual signature in propagated calls (manuscript 1)

Here I conducted sound transmission experiments and performed acoustic analyses to characterize and quantify the degradation of the individual signature of female and male zebra finch distance calls during propagation. The aim of this descriptive analysis was firstly to determine the acoustical features that are the most informative about individual identity at different distances; secondly to test the discrimination performance of a set of "templates", which determined different levels prior information, in a theoretical framework for auditory processing; and thirdly to investigate the potential differences in the individual signature of female vs. male calls in the way they degrade with propagation. In this study, I also contributed to the development of a novel approach for classifying sounds by using a linear discriminant analysis that operated directly on a full and invertible spectrographic representation of the sounds. This data-driven approach was compared to the more traditional approaches that have been used by bio-acousticians, where a small number of specific acoustical parameters (e.g. mean frequency, duration) are chosen to analyze and discriminate sounds.

2. Behavioral experiments testing the discrimination and learning abilities of female zebra finches (manuscript 2)

In this study, I first explored the reliability of mate recognition by females at a range of propagation distances by assessing their preference for their mate's calls, using an operant choice apparatus. Then, to assess the role of experience and to further distinguish the discrimination process from the recognition process, I conducted forced-choice conditioning experiments and compared the results of two different protocols. In both protocols, the females were asked to discriminate between the propagated calls of two unfamiliar males. In the first protocol, the females learned to discriminate the 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 females 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.

3. Electrophysiological experiments examining the neural substrate for individual discrimination of degraded calls (manuscript 3)

Here I tested zebra finches' neurons on their ability to discriminate not simply individual calls but ensembles of call exemplars produced by the same individual, and ascertained whether we could find neurons that could maintain their discrimination performance for this individual signature with increasingly degraded calls. My goals were to assess if such a capacity existed at the neuronal level, without prior familiarization or learning, and if so, to disentangle the effect on neural discrimination due to a decrease of SNR from effects due to sound intensity decreases. I also investigated the effect of sex on the between-voices discrimination performance, either at the level of the stimulus (male or female call) or at the level of the tested subject. To study these questions, I performed extracellular electrophysiological recordings on anesthetized zebra finches, using multi-channeled electrode arrays and taking distance calls with different levels of propagation-induced degradation as sound stimuli.

In the course of this research project I shared my time between the University of Saint-Etienne, France, and the University of California Berkeley, USA. The sound transmission experiments and part of the behavioral tests were conducted in France, and the electrophysiological experiments and the other part of the behavioral tests were performed in California.

MATERIALS AND METHODS

II. MATERIALS AND METHODS

A. Animal subjects

1. Housing conditions

a. In the ENES laboratory (Saint-Etienne, France)

The subjects (*n* = 10 adult male-female pairs, used for **manuscript 2**) were raised in the ENES laboratory aviary and kept in a temperature-controlled room (between 23 and 25°C) on a 14L/10D photoperiod with adapted wavelengths. Food (seed mix and egg mix) and water were given *ad libitum*, the birds had access to cuttlebones at all times and were supplemented once a week with fresh lettuce, and with vitamins if needed. Pairs were housed in separate cages (38 cm W x 24 cm D x 40 cm H) in the same room, having visual and vocal contact with each other.

b. In the Theunissen laboratory (Berkeley, California)

The subjects (n = 7 adult females for **manuscript 2** and n = 8 adults, 4 males and 4 females, for **article 3**) were raised in UC Berkeley's animal facilities and kept in single-sex cages (56 cm W x 35 cm D x 40 cm H), having visual and vocal contact with all other birds in the colony. The room was temperature-controlled (between 22 and 24 °C) and on a 12L/12D photoperiod. Food (seed mix) and water were given *ad libitum*, with a supplementation of lettuce and egg mix once a week.

2. Isolation of the zebra finch mated pairs

These pairs were used in **manuscript 2**. Prior to the experiments, I observed the birds of the ENES lab colony, housed in a mixed-sex room of about 40 birds. I sat outside of the room, behind a two-way glass, so as to observe their undisturbed behavior, and could identify each bird with its colored leg bands. Nests and nesting material (hay, cotton) had been installed in the rooms so as to encourage reproductive behavior. I selected for my experiment pairs for which I had seen repeated behaviors such as allopreening, nest building and guarding, and eggs incubation. I only chose partners that showed a strong and steady bond over the 2 to 3 weeks of observation. I isolated 20 pairs for the duration of the behavioral experiments. It is to be noted that due to various reasons (inability to record distance calls from the male, unexpected deaths, bird unable to perform the operant task), only 10 females from these 20 pairs were tested successfully, thus the number of 10 subjects given in **manuscript 2**.

One should note that the birds in both aviaries (ENES lab and Theunissen lab) were kept on different photoperiods, and I chose to keep them on the same day/night cycle during the experiments so as to avoid disturbing them.

B. Distance calls recordings and sound transmission experiments

Creating the database of propagated calls was the cornerstone on which all experiments and acoustical analyses were conducted. Here I describe the recording of the 'clean' distances calls that were used to record the propagated calls, and the "fieldwork", a.k.a. the sound transmission experiments.

1. Recordings of distance calls

For my thesis work I used two different distance calls databases, both recorded using birds from the ENES colony. The first database was recorded by Julie E. Elie during her PhD using unpaired zebra finches, and was used for all three articles. I recorded the second database that was used for **manuscript 2** using the mated pairs that I isolated as explained previously (II.A.2). The recording of the mated pairs database took place in two sound-proof rooms connected by a small soundproof corridor; the male and female of each pair were kept in two separate cages and placed on the furthest end of each room, 50 cm high from the ground, and about 4 m away from each other. The birds could hear but not see each other, so as to promote vocal interactions between the partners. Both birds were recorded using a microphone (Sennheiser MD-42) placed 0.2 m above the cage (*Fig. 6*) and connected to a Marantz Professional Solid state recorder (PMD 670; sampling frequency: 44100 Hz). Conditions of temperature, food and water availability were the same as in the aviary. For the behavioral experiment described in **manuscript 2**, I isolated 10 distance calls from each male and normalized them by matching the maximum values of the sound pressure waveforms. These calls were used to create the "mated males" propagated calls database (see below).

Fig. 6: Recording of distance calls.

The recording of the first database performed by Julie Elie took place in the same rooms, with the same equipment but using unpaired males and females. Each bird was recorded individually, in the presence of two females placed 3 m away and used as an audience to minimize stress. The recorded bird was stimulated with distance calls playbacks from previously recorded conspecific birds. From this database I retrieved 16 calls from 32 individuals (16 males and 16 females), and these calls (256 total for each sex) were used to create the "unpaired birds" propagated calls database. The calls were normalized as explained above.

2. Sound transmission experiments

We recorded the propagated calls for the unpaired birds database on a flat field (*Fig. 7*) near Bellegarde-en-Forez (Loire, France) on the 3rd of October 2010 in the afternoon (cloudy weather; no wind; temperature = 11°C). The propagated calls for the mated males database were recorded on the 1st of March 2011 around noon, on the same field, (cloudy weather; wind < 5 km/h, temperature = 10°C).

Fig. 7: Recording of propagated distance calls.

We followed the same procedure for both recordings, as follows. The complete calls database was broadcasted from a Marantz Professional Solid state recorder/player (PMD671) connected to a MegaVox speaker (PB-35W, Anchor Audio Inc.) placed 1.3 m high so as to avoid excessive ground reflection interference. The speaker volume was set to obtain a sound level of 70 dB SPL at 1 m (Velleman Sound Level Meter DVM 1326) to match typical levels of the natural distance call in the zebra finch (Vignal et al., 2008). The sounds were recorded with a Schoeps microphone (MK4 cardioid, on a CMC6-U base) equipped with a Schoeps Basket-type Windscreen (W20) and set 1.3 m high. The microphone was connected to a second Marantz recorder/player (PMD671; sampling frequency: 44100 Hz). All the calls of each individual bird were dispatched along the audio sequence, in order to avoid any context effect (e.g. changes in the background noise). We recorded the calls sequence 1 m (for calibration), 2 m, 16 m, 64 m, 128 m and 256 m away from the source, three to four times for each distance for the unpaired birds database, and 16 m, 64 m and 256 m away, twice for each distance, for the mated males database.

3. Extraction and pre-processing of the propagated calls

The main idea in this pre-processing phase was to retrieve each individual call from the recorded sequences and assign it the right bird individual and call number. As identifying the degraded calls directly on the propagated sequences would be impossible to achieve with enough certainty and precision (especially at long distances), I used the original sequence (for which the bird and call numbers for each individual call were known) and cross-correlated the propagated recordings at each distance with the original sequence using a custom-made script on Matlab (Mathworks). I then cut the whole sequence files into unique distance calls files, named after the bird and call numbers, and the propagation distance.

Next, I compared the signals from the two to four recording sessions (depending on the database) for each call and at each propagation distance, by ear or by a visual assessment of spectrograms (using the Praat sound analysis software; Boersma, 2002)) when necessary, to select signals that hadn't been impaired by an unexpected sharp noise that was not relevant to the propagation study (eg. birds calling in the vicinity, farm animals calls or human-related activity).

On R (The R Project for Statistical Computing; Team, 2005)), I then used the timer function from the Seewave package (Sueur et al., 2008) on the calls that had been cut from the original sequence to find the exact coordinates of the beginning and ending of each call in the audio files. As the files from all recorded calls had been cut identically after being cross-correlated to the original sequence, I ensured that these coordinates matched exactly the beginning and end of each call in the propagated audio files (*Fig. 8*). I then cut the sound files to the length of the longest call for males and females separately (longest call duration for mated males: 0.25 s, for unpaired males: 0.29 s, and for unpaired females: 0.44 s), and used the coordinates to ensure that each call was centered inside this window.

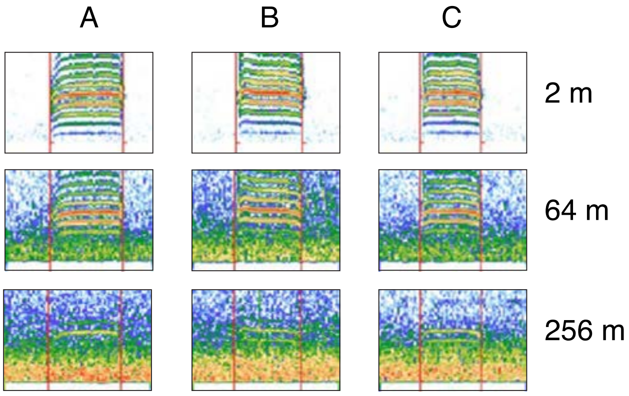


Fig. 8: Coordinates for the beginning and ending of 3 calls (A, B and C) at 3 propagation distances. Note that the color scale for the spectrograms is different for each distance.

C. Analysis of propagated calls

There are different levels of sound analysis in this work. First, I will present various representations of acoustic signals that are useful for visualizing different spectral and temporal aspects of the sounds, and calculating parameters describing them. I will then explain the issues that I encountered in dealing with degraded signals to characterize in a systematic way the variously degraded vocalizations. Finally, I will detail the acoustic analysis that was conducted on the call database (**manuscript 1**).

1. Representations of an acoustic signal

a. Oscillogram

The oscillogram is the graphical representation of the sound wave that is closest to the physical phenomenon at the origin of a sound: the air pressure fluctuations are represented as a function of time (x-axis), and the rise and fall of pressure translates into the sound amplitude (y-axis) (*Fig. 9*). While the oscillogram is useful to measure temporal variables such as the duration of a call, information about the spectral content of the call is not directly available in this representation. Thus, the oscillogram can be defined as a description of the sound in the "time domain". While a pure tone (characterized by a unique frequency) would be represented as a perfect sinusoid on the oscillogram, more complex sounds such as animal vocalizations consist of many frequency components that vary in magnitude (i.e., amplitude) and phase (relative timing of each frequency component). The oscillogram only shows the additive result of these frequency components.

b. Spectrogram

In a spectrogram (also called sonogram), the sound energy is plotted as a function of time (x-axis) and frequency (y-axis) (*Fig. 9*). This time-frequency representation can be obtained from the time-amplitude representation of the sound (i.e., the oscillogram) by calculating the Short-Time-Fourier-Transform (STFT). I will not go into the mathematical details of this operation, but in intuitive terms, the Fourier Transform (FT) breaks down the waveform into sinusoids of different frequencies, that is, breaks down the sound wave into a set of frequency components, characterizing their amplitude and relative phase. It is a beautiful thing indeed that a continuous waveform can be divided into a set of pure sine waves, and that re-combining these sine waves at their given amplitude and relative phase will reconstitute the original signal exactly as it was. On a more pragmatic note, the phase information is not explicitly shown on a spectrogram but it can be recovered in the spectrogram inversion.

In most sounds emitted by animals, frequency and amplitude vary over time. In order to describe these variations accurately, the STFT uses a fixed temporal window that slides down the signal's time dimension, obtaining a set of amplitudes as a function of frequency for each time window. The resolution of the spectrogram thus depends on the size of the temporal window: taking a large window would give precise information about all the frequency components (i.e., about the frequency domain) but little information about how the signal varies in time (i.e., about the time domain). Conversely, using a small window would precisely describe the time domain but would result in coarse frequency resolution. One can now understand that it is not possible to have high resolutions in the temporal and spectral domains at the same time; choosing the window size is therefore a tradeoff, and the choice depends on the type of signal that is analyzed and the purpose of the analysis. For the acoustic analysis of degraded sounds (**manuscript 1**) I chose the temporal window that gave the best results in terms of discrimination of individual identity (see Results): it was a Gaussian window (symmetric in the temporal and spectral domains) of 2.27 ms, corresponding to 70 Hz.

The spectrogram enables to visualize the spectro-temporal structure of an acoustic signal. The distance call of zebra finches (as most animal vocalizations) consists of a harmonic series (i.e., harmonic bands that are multiples of a fundamental frequency). In the spectrogram of a female call represented in *Figure 9*, the fundamental frequency is barely visible, and the harmonic bands with the most energy (as represented by redder colors in the color scale) are found between 2 and 4 kHz.

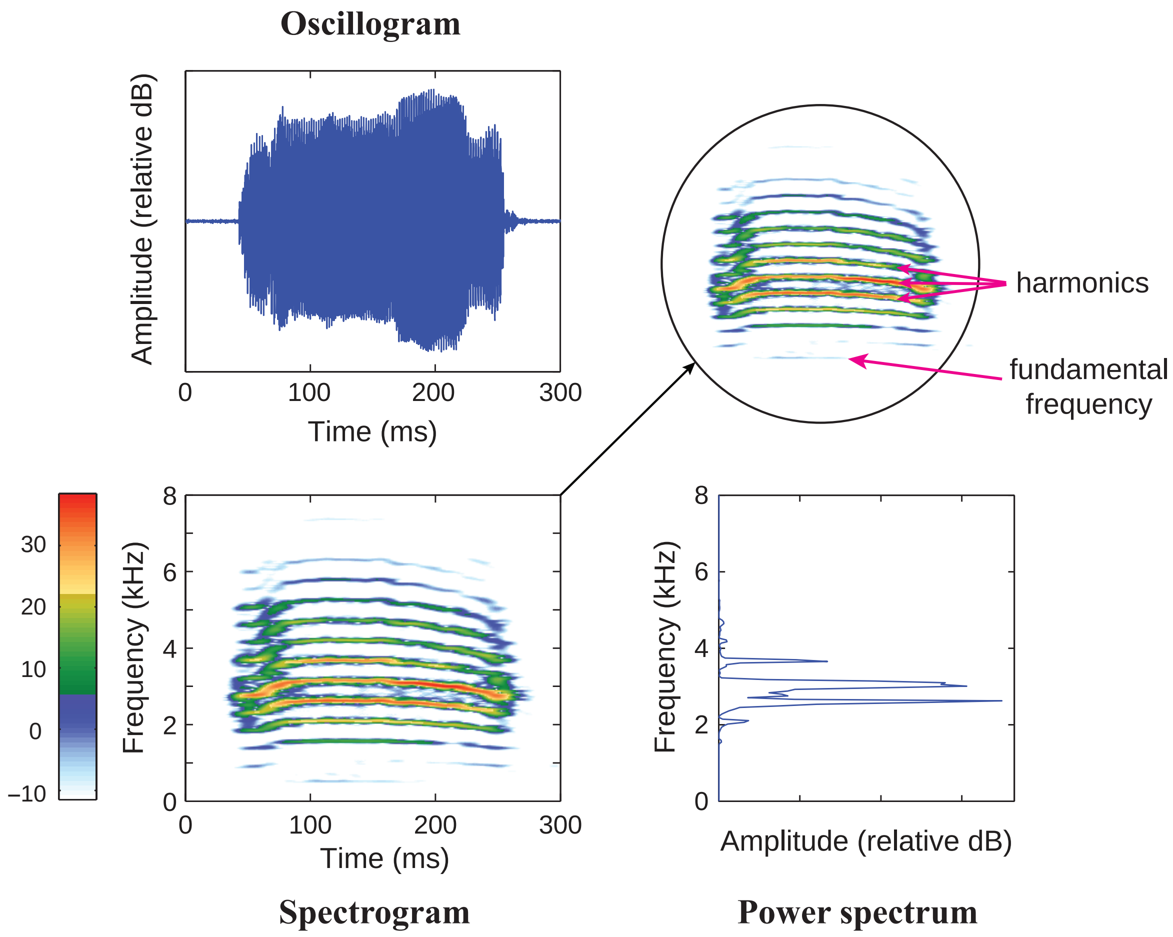


Fig. 9. Three acoustic representations of the same female call. Top left, the oscillogram shows the signal's amplitude as a function of time. Bottom left, the spectrogram is a representation of the sound's energy as a function of time (x-axis) and frequency (y-axis). The color scale is in relative dB as shown on the color bar. Bottom right, the power spectrum represents the power of the sound as a function of frequency. Finally, harmonic structure of the call is shown on the top right of the figure, using the spectrogram.

c. Power spectrum

The power spectrum, or frequency spectrum, is the amplitude square (the power) of the sound represented as a function of frequency. It is a useful representation of the distribution of the sound energy over the frequency bandwidth of the sound. For calls showing a low frequency modulation over time (as is the case for the female call showed in *Figure 9*), peaks of energy corresponding to the most powerful harmonic bands are clearly visible on the power spectrum.

In my studies (**manuscript 2**), I have used the power spectra to quantify the similarity between sounds by calculating the correlation between the mean power spectra of two individuals. This basic spectral analysis served to assess the difficulty for subjects to discriminate between calls depending on their similarity.

d. Modulation power spectrum

The modulation power spectrum (MPS) is a second-order representation derived from the spectrogram, which enables to visualize and quantify the joint spectro-temporal modulations of calls. It has been developed by Singh and Theunissen (2003), and the mathematical calculations for the joint second order statistics of the amplitude envelopes along the temporal and spectral dimensions are presented in their paper. Here I will explain how to 'read' a MPS, as this representation, while being useful to visualize higher-order structures in a signal, is not very intuitive. As shown on *Figure 10*, temporal modulations are shown on the x-axis and characterize the power in the temporal amplitude envelope as a function of temporal frequencies (in Hz). For example, signals that contain fast changes in their amplitude envelope (such as a quick onset) are characterized by power at the higher temporal modulation frequencies. The spectral modulations are shown on the y-axis and characterize the power in the spectral envelope in terms of spectral frequencies (in cycles per kHz). For example in the female call, the fundamental frequency is around 500 Hz, which corresponds to 2 cycles per kHz; the pitch being steady, this shows on the MPS as a very visible area of energy on the y-axis at 2 cyc/kHz. Energy at lower values on the y-axis indicate lower spectral modulation patterns, such as differential energy in odds vs. even harmonics (for a fundamental frequency at 500 Hz, enhancing every second harmonic represents a spectral modulation at 1 cycle per kHz), or formants created through vocal tract filtering. In the male call, the higher fundamental frequency yields a lower spectral modulation rate, and the unsteadiness of the pitch due to the important frequency modulation shows in the 'smeared' energy area around the pitch zone. Background noise, which produces incoherent modulations in the spectral and temporal domains, is shown around the zero values of both axes.

Any point away from these axes represents joint spectro-temporal modulations such as upsweeps (in the area with negative temporal modulation) or downsweeps (in the area with positive temporal modulation). On the examples shown in *Figure 10*, the upsweep and downsweep patterns at the beginning and end of the female calls are shown through a spread of energy at 2 cyc/kHz along the horizontal dimension. For the male call, the important downsweep component is clearly represented in the energy area found at high temporal modulation rates, compared to the less prominent upsweep component at the beginning of the call.

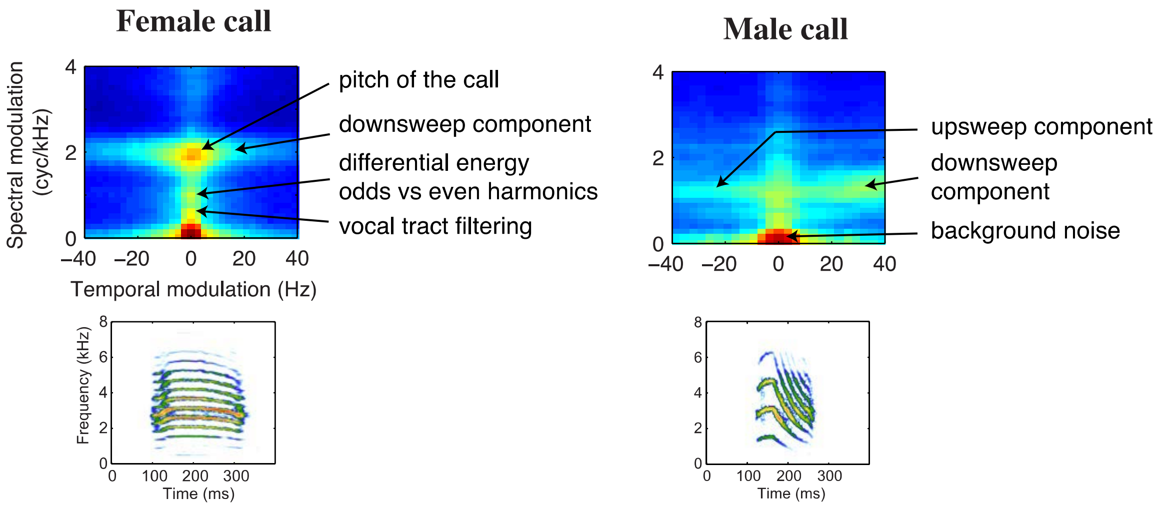


Fig. 10: Modulation power spectrum (MPS) of a female and male distance call. Temporal modulations (in Hz) are shown on the x-axis, and spectral modulations (in cycles/kHz) on the y-axis. The spectrogram of the corresponding call is shown below each MPS.

2. Choice of the parameters used for the acoustic analysis

When studying 'clean' signals, recorded in laboratory conditions for example, it is possible to measure very precise characteristics of vocalizations, such as the signal's duration or parameters closely describing the fundamental frequency and its variations in time (Vignal et al., 2008). However, in propagated calls, the calculation of such parameters with reasonable precision is made impossible because of the sound degradation: the fundamental frequency usually disappears in the background noise (*Fig. 8*), and the low resolution of the temporal characteristics of the signal renders the calculation of even basic measurements, such as the call's duration, difficult.

In order to circumvent these issues, and to be able to calculate acoustic parameters in a systematic way, regardless of propagation distance, I first used a set of parameters that did not rely on such precise measurements but rather separately described the amplitude in the spectral domain (using the spectral envelope) and in the temporal domain (using the temporal envelope). Each envelope was converted to a density function on which five parameters were calculated (*Fig. 11*): the mean, standard deviation, skewness (a measure of the asymmetry of the envelopes shape), kurtosis (a measure of the peakedness in the shape of the envelopes) and entropy (which captured the overall variability in the envelopes; see Results). I also calculated a second set of parameters taking into account the entire spectrogram of the calls (*Fig. 11*). Because of the high dimensionality of such a representation, I used a Principal Component Analysis (PCA) to reduce the dimensionality to a manageable set of parameters (the principal components) that described the variability in the calls. This second set of parameters took into account joint spectro-temporal modulations in the signals, which proved very useful in the statistical analysis that followed. Finally, in order to calculate these parameters using the same time frame for all calls, I set the temporal window to the length of the longest call (separately for each sex), and centered all the calls inside this window.

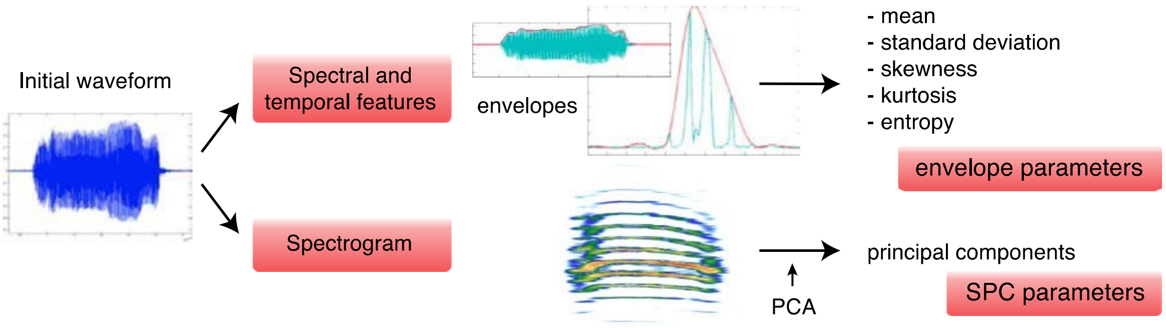


Fig. 11: Procedure for extracting both sets of parameters. The envelope parameters are shown on top, and the parameters obtained from the whole spectrogram at the bottom.

3. Discriminant function analysis

To test for the presence of an acoustic signature in the propagated calls (**manuscript 1**) I used a multivariate method, the cross-validated discriminant function analysis (DFA), taking into account a set of parameters describing the calls at various distances, as explained above. The DFA finds the linear combinations of the parameters that maximally separate the data points (the calls) that belong to different categories (the bird identities), that is, clusters the calls of each individual so that overlap between individuals is minimal. For each DF, the value for each parameter is multiplied by a weighting coefficient (the array of these coefficient corresponding to each parameter is called an eigen vector). Each DF is also assigned a coefficient (the eigen value) that accounts for the amount of variability in the data that is explained by the DF. The set of significant DFs define a Euclidian subspace where the calls can be projected, and distances between calls can be calculated.

The cross-validation consists of first calculating a set of discriminant functions using a fitting dataset (*Fig. 12*), and then testing these functions on a validating dataset (non-overlapping with the fitting dataset) by assessing the number of calls that have been correctly classified for each individual and calculating a percentage of correct classification. In the example shown on

Figure 12, for each individual, all calls but one are used to calculate the DFs in the fitting dataset. The remaining call is used as part of the validating dataset to test for its classification as belonging to the correct individual: each call from this validating dataset is projected on the Euclidian subspace defined by the DFs, and its distance to the clusters of calls pertaining to each individual is calculated. The cluster yielding the smallest distance designates the individual to which the call is attributed. This process is repeated a number of times, randomly selecting fitting and validating datasets, and the classification results are incremented in a confusion matrix representing the joint probabilities of the predicted vs. actual individuals. A percentage of correct classification can be calculated from this matrix, as a measure of discrimination performance.

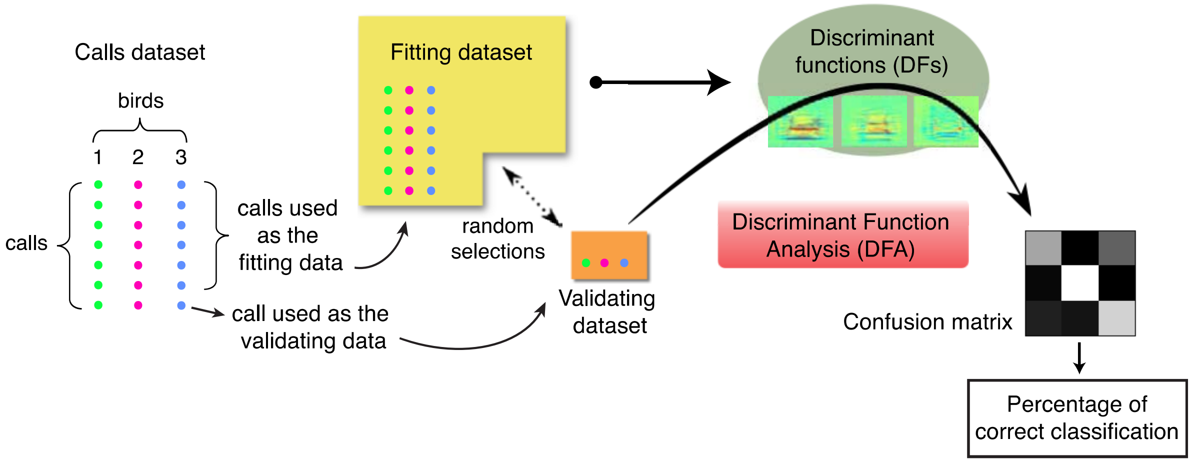


Fig. 12: Cross-validated discriminant function analysis (DFA). The different steps of the procedure are illustrated (see main text), with an example database consisting of 8 calls from 3 individuals.

D. Behavioral tests

In the behavioral experiments testing the discrimination abilities of female zebra finches presented in **manuscript 2**, I used two different operant conditioning setups: first, I tested females in a preference test where their choice was unconstrained; and second, I used a forced-choice discrimination task for which the call yielding the reward was imposed. While the first experiment tested females in the context of mate recognition, the second test was aimed at investigating the limits of their discrimination abilities.

1. Preference test

In this experiment I tested the preference of isolated females (n = 10) for hearing the call of their mate compared to the call of a familiar male, both calls being propagated at the same distance. The setup consisted of an experimental cage with a central body (30 cm W x 34 cm D x 34 cm H) where food and water were available *ad libitum* and that contained a single perch (*Fig. 13*). On each side of the cage, an opening (10 cm W x 10 cm H) led to a side arm (20 cm W x 10 cm D x 26 cm H) containing a perch and was equipped with infrared sensors that monitored when the bird entered the arm. I used a custom-made software to monitor the subject's activity on the perches and trigger playbacks as follows: a hop on a side perch broke the infrared beam and triggered the playback of a call from a loudspeaker (Bravo Allroom, Audio Pro, Sweden) placed 20 cm away from the same side arm. Depending on the side arm, this call was randomly selected either from the 10 calls available for the tested female's mate or from the 10 calls of a familiar male. Sound stimuli were broadcasted by either of the two loudspeakers connected to an amplifier (Yamaha Natural Sound Stereo Amplifier, AX-396, Hamamatsu, Shizuoka, Japan) and a laptop. I calibrated the intensity of the sound stimuli by setting it at 70 dB SPL for the sounds recorded at 1 m (typical level of a natural distance call) and used that gain setting for all playbacks, ensuring that signals propagated at further distances were emitted at the lower intensity level that matched the amplitude loss due to natural propagation. Each subject was housed alone in the experimental cage, which was placed in a sound attenuation chamber (internal dimensions: 1.8 m W x 1.4 m D x 2.2 m H; Silence-Box, Tip-Top Wood, Saint-Etienne, France).

Each female zebra finch was subjected to 3 trials, with calls recorded at 16 m, 64 m and 256 m, in random order. Each trial consisted in two sessions and started with a habituation period, enabling the subject to get used to the setup and learn which side arm was associated with which individual's calls (Mate or familiar male). I put the subject inside the cage on Day 1 at 1800h, and placed small food dispensers on each side arm, to encourage the subject to explore the arms and hop on the side perches; at that time however, the amplifier was turned off so that no call could be heard from the speakers. On Day 2, the amplifier was turned on and the subject could hear the playback when hopping on either side perch, the mate's calls (Mate) being triggered on one side (e.g., left side) and the familiar male's calls (Non-Mate) on the opposite side. At 1030h on the same day, I entered the chamber, removed the food trays on either side arms and placed a larger food dispenser in the middle of the front pane in the central part of the cage (*Fig. 13*). After letting the bird settle down for 30 min after my intrusion, the data collection of the subject's activity on each side perch started at 1100h, the side assignment for the mate and familiar male's calls being unchanged (session 1). At 1400h on Day 3, the side assigned for each male's calls was swapped (e.g., mate's calls now on the right side), and the second session started, so as to balance the experiment and control for possible side preferences of the subject. The amplifier was turned off during night time (from 2200h to 0800h), so that the subject could not hear any playback when hopping on the side perches. The trial was ended at 1700h on Day 4, which insured that the subject's activity was recorded during the same amount of time (17 h) for each Mate/Non-Mate side assignment. Using the perching events on the side arms, I then analyzed the data to assess if the side assignment for the mate and the distance had an influence on the females' choice to perch on either side arm (see **manuscript 2**).

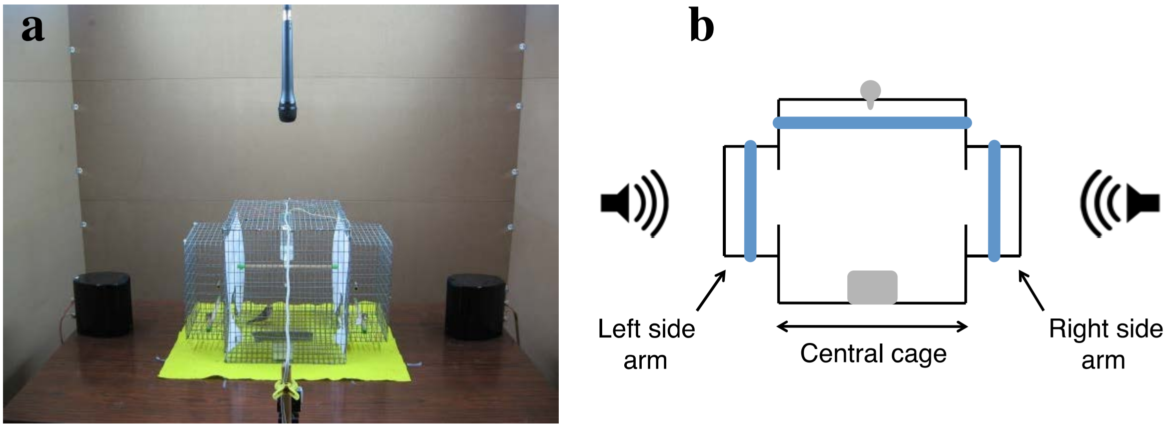


Fig. 13: Preference test apparatus. (**a**): picture of the setup inside the chamber; (**b**): diagram representing the experimental cage (as set during the recording sessions). Perches are represented in blue, food and water dispensers in light gray. By hopping on the perch in each side arm, the female subject triggered the playback of either her Mate's distance call or a familiar male's distance call (Non-Mate) from the speaker placed beside the side arm.

2. Discrimination task

The purpose of this experiment was to further investigate the ability of females to discriminate between two degraded vocal signatures, without relying on the female's preference for her mate as an indication that discrimination occurred. Using a pecking key apparatus and a forced choice operant procedure, I tested fasted females to discriminate a rewarded stimulus from an unrewarded stimulus, the reward being access to food for 10 s. The subjects could trigger the playback of calls at will by pecking on a key. At any time they could choose to attend the full duration of the stimulus or peck again to interrupt the current stimulus and trigger the next one. The access to the feeder was only permitted when the bird chose to fully attend to the rewarded (Re) stimuli. The discrimination task apparatus consisted of a modular test chamber (interior dimensions 31 cm x 24 cm x 29 cm; Med Associates Inc, St. Albans, VT, USA) placed in a soundproof booth (Acoustic Systems, MSR West, Louisville, CO, USA; interior dimensions 76 cm x 61 cm x 49 cm). The experimental panel consisted of a pecking key placed 20.5 cm high from the floor and accessible through a wooden perch (*Fig. 14*). Below, a feeder containing seeds could be made accessible or not to the subject, depending on its appropriate response to the playback. Each acoustic stimulus used for the playback consisted of a sequence of 6 distance calls randomly selected from the 16 available calls of the same male individual for the same distance in the "unpaired birds" propagated calls database, and randomly distributed within a 6 s window. Acoustic stimuli were broadcasted by a computer connected to an amplifier (Technics, Matsushita Electronics SA-EX140, Osaka, Japan) and a loudspeaker (Bose model 141, Framingham, MA, USA) placed 20 cm from the test chamber (sound level calibrated as in the preference test to match the natural intensity levels at each propagation distance). The computer was also connected to the test chamber apparatus to record pecking events, play sounds and activate the feeder in real-time with a single customized program written in Matlab©.

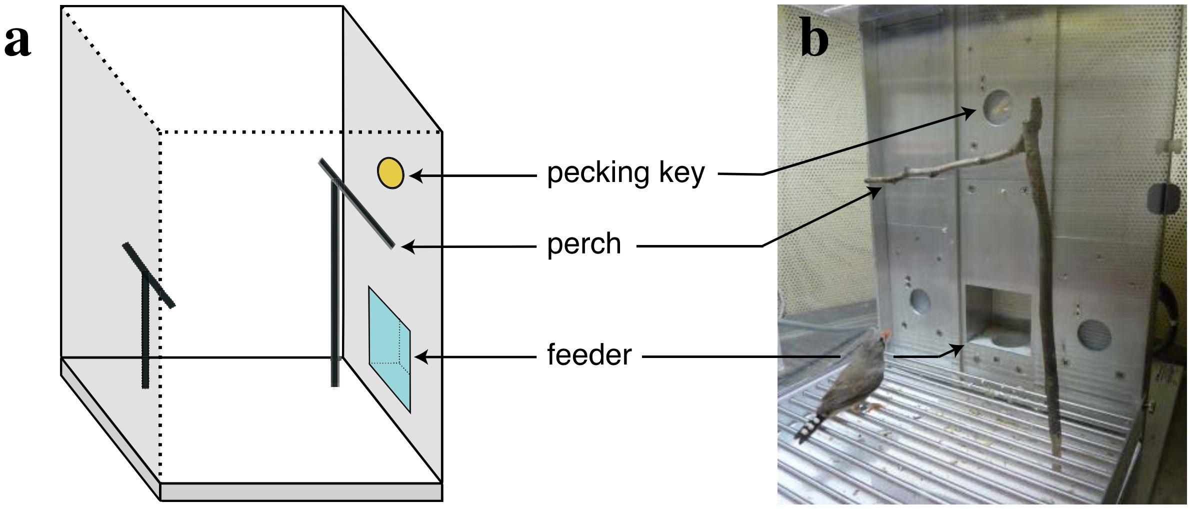


Fig. 14: Discrimination task apparatus. (**a**): diagram of the pecking test chamber; (**b**): picture of the apparatus. Subjects pecked the key to trigger the playback of distance calls from either the rewarded or the unrewarded male; appropriate response to the playback ensured the subjects a greater access to food. The water dispenser is not represented in this figure.

The conditioning procedure consisted of 4 tests conducted for 4 consecutive days. One test consisted of three 30 min trials separated by two 90 min-long rest periods. The pecking key's light was used to distinguish the trial period (pecking light on) from the rest period (pecking light off). The 30 min countdown for each trial started when the subject pecked the key for the first time. When pecking the key during a trial (*Fig. 15*), the female triggered the playback of calls from either of two males used as stimuli: the Re male (with a probability of 0.2) or the NoRe male (with a probability of 0.8). She could then go to the feeder and wait until the end of the 6 s playback to get a reward for the Re stimuli, or interrupt it by pecking again to trigger a new stimulus. Because the time windows for pecking (the three 30 minutes trials) were limited and most stimuli were NoRe, the subjects were motivated to interrupt the NoRe stimuli until they obtained a Re stimulus, at which point waiting until the end of the playback would ensure them access to seeds for 10 s. Interrupting the playback of a Re stimulus eliminated the possibility of reward following this playback. The subjects were thus tested on their ability to interrupt the Non Rewarded (NoRe) stimuli and refrain from interrupting the Rewarded (Re) stimuli, this behavior ensuring them a greater access to food.

To motivate the subjects to use the pecking key for food reward, I fasted them 20 hours prior to the beginning of the experiment and maintained them in a fasted state (85-90% of their free feeding weight) for the whole experiment by only giving them 1.5 g of seeds per individual after each daily test. Every day, I monitored the subjects' weight before starting the test. As approved by the Animal Care and Use Committee of UC Berkeley, the criterion for interrupting the fast was a loss of weight superior to 15% of the initial weight recorded before the fasting started. No bird was taken out of the procedure following this criterion.

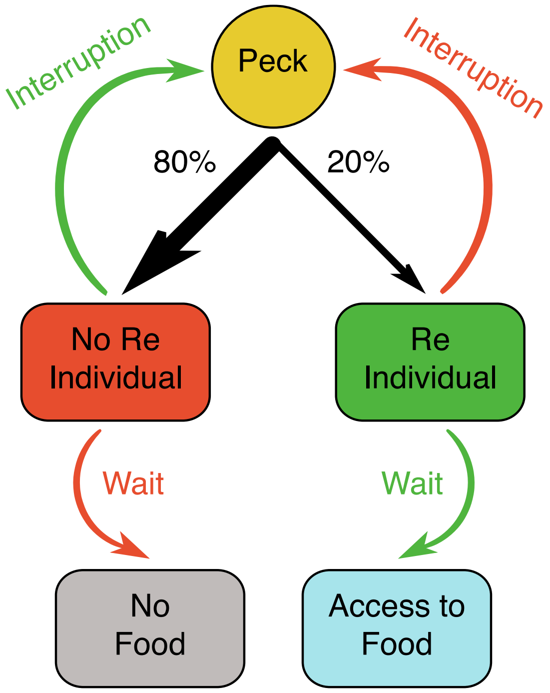


Fig. 15: Experimental protocol for the discrimination task. Straight black arrows represent the playback options resulting from a peck on the key (with the probability for the occurrence of each option), and curved arrows represent the behavioral options for the subjects on hearing the playback ('correct answers' in green, 'wrong answers' in red).

Prior to the experiments, the subjects were shaped when they first started using the setup. The 7 females that I used for the experiment presented in **manuscript 2** had already been shaped by Julie Elie. She had also tested them in discrimination tasks, using stimuli that were different from the ones I used, and non-propagated. This shaping lasted for two days. On the first day the bird would acclimate to the test chamber, find the feeder and get access to seeds; then, the bird ran a long trial (1 to 3 hours) where it learned that pecking the key triggered a vocalization, and that one of the two vocalizations was followed by a timed access to food. The stimuli used for the shaping were two 6 s sequences of songs recorded from 2 different males in the colony. To make it easier for the subject to learn, the probability to get a Re stimulus when pecking was set between 0.5 and 0.9, depending on the subject's motivation. On the second day, 3 regular trials were conducted with a probability of triggering the Re stimulus of 0.2 (as in the experimental test) and the subject learned to interrupt the stimuli. A female was considered ready to run experiments when it interrupted at least 20% of the NoRe stimuli and the difference in percentage of interruption between the Re and NoRe stimuli was at least 20%.

For every subject that I used, the experiment started on Day 0 with a shaping test, using the same song stimuli used to train them to use the apparatus. This ensured that for every subject, who had last used the apparatus at different times prior to this experiment, the last experience they had with the setup before the actual test was the same.

E. Electrophysiology experiments

Each bird used as subject in the electrophysiological recordings underwent two surgeries, the first one to localize the auditory areas and prepare the bird for the recording day, and the second one on the recording day to place the electrode array in the brain. The electrophysiological recording in itself lasted for 7 to 10 hrs. The bird was then euthanized, and its brain was processed to histological purposes.

1. *Animal procedures*

For each of the 8 birds tested in **manuscript 3**, the protocol was conducted as follows. Two days before the neurophysiological recording, I anesthetized the subject with isoflurane (2

L/min for the inducement, then 0.9 L/min after stabilization) for the preparatory surgery. The aim of this surgery was to mark the emplacement of the auditory area (including CM, field L and NCM) on the skull and adhere a steel post on the subject's head that would be used during the electrophysiological recording to hold its head stationary. It consisted of the following steps: removing a small region of skin on the top of the skull after having immobilized the bird on a stereotaxic apparatus, using ear bars and a beak holder; removing a 1 mm by 2.5 mm rectangle of the top layer of the skull around the auditory area zone; marking the reference for the most caudal and medial electrode (0.25 mm lateral and 0.25 mm rostral from the midsagittal y-sinus) with China ink to guide the electrode array penetration during the recording surgery; and finally, gluing a stainless steel post onto the skull with dental cement, cranially from the opening in the skull. Depending on the subject, this opening was either made on the left or right hemisphere, or both. After the surgery, the subject was allowed to recover for 2 days before undergoing the recording surgery that preceded the electrophysiological recording. The main reason for performing this preparatory surgery on a separate day from the actual recording was due to the fact that the most precise way of localizing the auditory zone from the y-sinus reference (*Fig. 16*) was to hold the head of the bird in a precise and standardized position, using a beak holder and ear bars. However, obstructing the bird's ears is not advisable when recording responses to auditory stimuli, which is why the stainless steel post was fixed on the bird's skull in order to hold its heard while keeping its ears free during the recording.

On the recording day, I fasted the bird for 1 h then anesthetized it with three intramuscular injections of 20% urethane (30, 25 and 20 μL) administered at 0.5 h intervals. The bird was then placed in the stereotax apparatus and its head immobilized by attaching the steel post to the stereotax frame. I removed the lower layer of the skull and dura mater in the rectangle opening that I had made previously, and using a microdrive I lowered into the brain an extracellular 16-channel electrode array (Omn1010, Tucker-David Technologies Inc., Alachua, FL, USA), consisting of 2 rows of 8 electrodes (width 0.5 mm, length 2 mm, distance between two electrodes within one row: 250 μm; distance between rows: 500 μm) (*Fig. 16*). Each electrode had previously been coated with DiI stain (Invitrogen, Eugene, Oregon, USA) so as to facilitate the electrodes localization in the brain during the histological analysis.

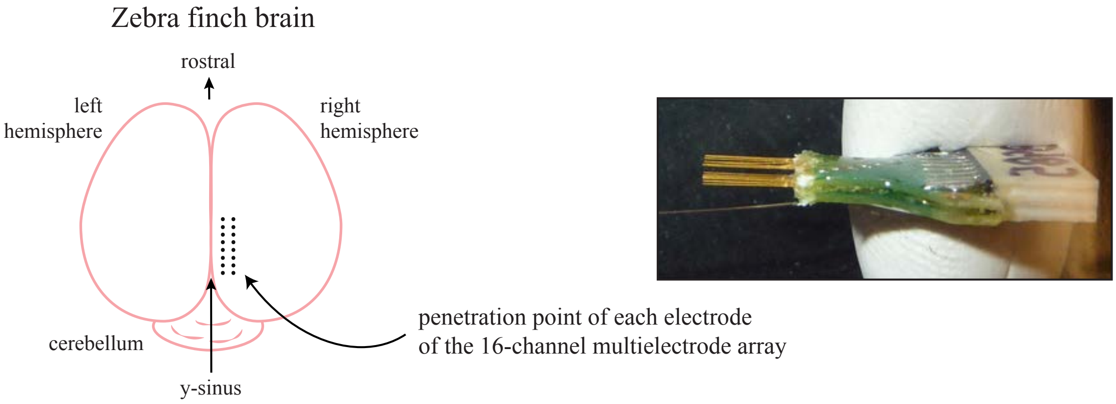


Fig. 16: Electrode penetration area in the brain for an electrophysiological recording in the auditory area. I recorded neurons in right hemisphere (as shown on the figure), left hemisphere, or both, depending on the subject. A picture of a multi-electrode array is shown on the right.

2. Electrophysiology recordings

For the recording, I placed the bird and stereotax in a double-walled anechoic chamber (Acoustic Systems, MSR West, Louisville, CO, USA) where a loudspeaker (Blaupunkt T-line, Berlin, Germany) was used to broadcast the stimuli. The volume of the loudspeaker was set to deliver zebra finch calls at 70 dB SPL (Digital Sound Level Meter, RadioShack, weighting type B) and was placed 20 cm in front of the subject's head. Using a microdrive, I then slowly lowered the electrode array further into the brain, down to 1-1.1 mm from the surface, and broadcasted a set of "search sounds" to the bird to test if the electrode array was recording neurons in an auditory area. These sounds were varied in their type (songs, calls, white noise) and were different from the stimuli used for the actual experiment. Using an integrated system from Tucker-Davis Technologies (TDT) recording extracellular voltages recorded for each electrode, I could hear and see the firing spikes, and visualize their waveforms. When the electrode array was deep enough to record auditory responses, I started the recording at this site. For each site, calls from one sex only were broadcasted, propagated at all distances (2, 16, 64, 128 and 256 m). The inter-stimulus interval was uniformly distributed between 1 to 3 s to prevent any rhythmic pattern that could potentially entrain the neurons or generate expectations. The presentation of all the stimuli was repeated 8 times and for each of these 8 trials the order of stimulus presentation was randomized. In this manner, any stimulus dependent adaption was avoided. When the recording was done for the site I pushed the electrode array further, at least 100 μm lower from the previous site, and proceeded to record a new set of stimuli at this new site.

3. Histological procedures

When all sites were recorded, I euthanized the bird by overdosing it with isoflurane and perfused it transcardially with phosphate buffered saline (PBS, Sigma Chemical co., St Louis, MO, USA), followed by 4% formaldehyde. I then removed the brain from the skull and sunk it in 4% formaldehyde, followed by 30% sucrose, before freezing it using liquid nitrogen. Next, I sliced the brain, frontally or parasagitally, in 20-μm-thick sections using a freezing microtome. I stained alternating brain sections with either cresyl violet or DAPI nucleic acid stain, in order to visualize the histological tissues (on the cresyl violet stained sections) as well as the electrode tracks (on the DAPI stained sections, with the DiI stain marking each electrode emplacement) across the brain sections. These observations were made using a Zeiss AxioImager M2 fluoresence microscope fitted with a camera (Retiga 1350 EX, QImaging).

Localization of the electrodes was a difficult process that involved measuring the distance from the entry of the electrodes to their deepest point and comparing it to the depth of the last recording site as shown on the microdrive used during the recording; recording site localization could then be achieved from the coordinates of each site obtained from the microdrive. Using well-known landmarks such as the *lamina mesopallialis* (LaM; known in the old nomenclature as hyperstrial lamina or LH) and the dorsal *lamina pallio-subpallialis* (LPS; previously called the medullary lamina or LMD) and differences in cell density as described in literature (Fortune and Margoliash, 1992), I then assigned recording sites to either CM (lateral: CLM, or medial: CMM), NCM, thalamo-recipient subdivision L2b or subregions L1 or L of the field L complex. I could not achieve further distinctions in the field L complex (L2a and L3) with enough certainty to assign recording sites as such, and these sites found in undefined areas were labeled Lx. Following guidelines found in literature (Vates et al., 1996), I chose to approximate the limit between CMM and CLM as being 800 μm away from the midline. I found this histological analysis very complex, owing to the fact that inferring 3-D representation on 2-D sections that were not always easily interpretable (due to folded or ripped tissue sections for example) was a difficult exercise, and that achieving a reasonable level of certainty in the site assignment was not always possible. The results for the histological analysis presented in **manuscript 3** are still preliminary at this stage, as I was able to localize only a small number of neurons.

4. Representation of neural responses

Neurons respond to stimuli that they are sensitive to by firing spikes, which are action potentials. These electrical impulses are essentially identical to one another (i.e., same amplitude and same width) for the same neuron. Thus, at the level of the single neuron, the coding of information will be a matter of how many spikes are emitted in response to a stimulus as well as their time-of-arrival; this action potential sequence as a function of time is called a spike train, and is a useful representation to visualize the temporal pattern of spike firing in response to a stimulus (*Fig. 17*). Another well-used representation is the peri-stimulus time histogram (PSTH), in which spike trains of all trials for the same stimulus are superimposed in time and used to construct a histogram (*Fig. 17*).

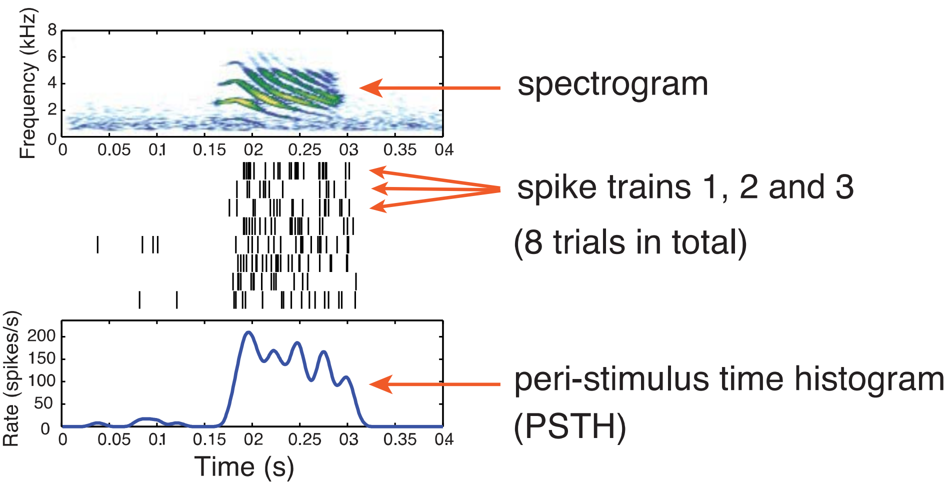


Fig. 17. Graphical representations of the responses of a single neuron to a stimulus. The spectrogram of the stimulus is shown on top. In the middle part, spike trains show the temporal sequence of spikes for each of the 8 trials (see II.E.2). At the bottom, the peri-stimulus time histogram (PSTH) is shown, representing the mean spike rate (in spikes/s) across all trials as a function of time.

RESULTS

Acoustic communication and sound degradation: How does the individual signature of zebra finch calls transmit over distance?

Solveig C. Mouterde[1,2*], **Frédéric E. Theunissen**[2], **Julie E. Elie**[2], **Clémentine Vignal**[1] **and Nicolas Mathevon**[1]

[1] Equipe de Neuro-Ethologie Sensorielle ENES-CNPS CNRS UMR8195, Université de Lyon/Saint-Etienne, 42023 Saint-Etienne, France

[2] Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720, USA

[*] Corresponding author: solveig.mouterde@gmail.com

Short title: "Individual signature in the zebra finch's call"

PLOS ONE, submitted

Question

How is the individual signature of zebra finches' calls degraded by propagation?

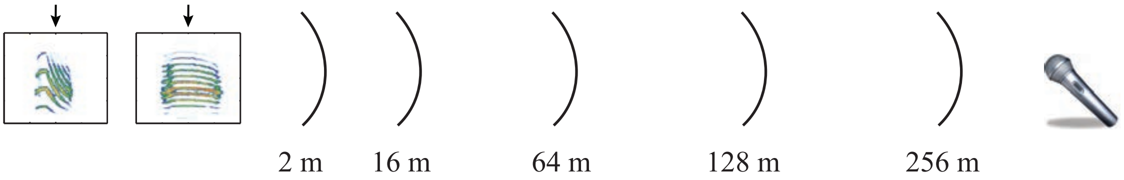
Methods

Sound transmission experiments

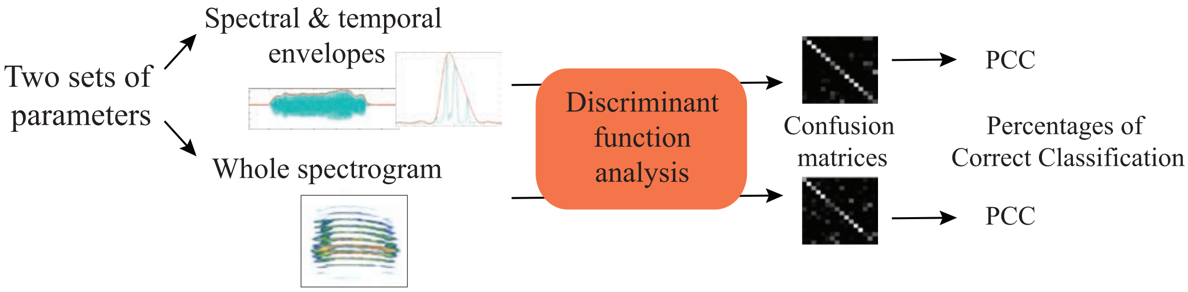
Distance calls of males and females...

propagated through natural environment...

...and recorded at various distances



Acoustic analysis



Two sets of parameters

Spectral & temporal envelopes

Whole spectrogram

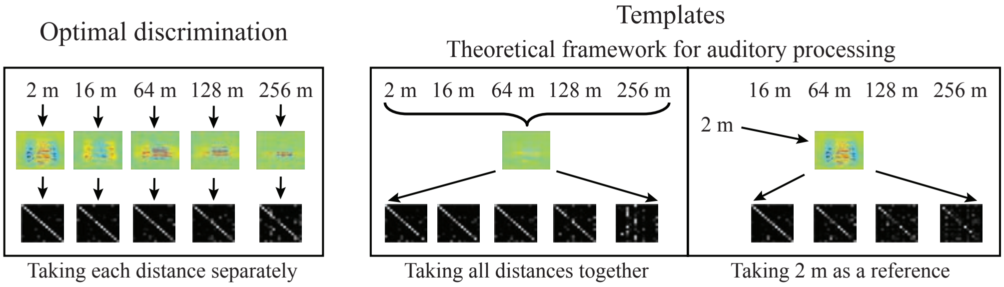
Discriminant function analysis

Confusion matrices

PCC

Percentages of Correct Classification

PCC



Three types of analysis

Optimal discrimination

Taking each distance separately

Templates

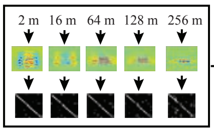
Theoretical framework for auditory processing

Taking all distances together

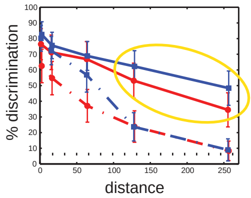
Taking 2 m as a reference

Results

Optimal discrimination



Individual signature is very resistant to long distance propagation...



...it is more resistant for males than females...

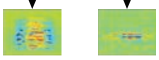
(this is due to the frequency modulation in male calls)

...and taking the whole spectrogram yields better discrimination than envelopes

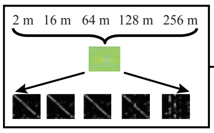
And also...

* information is carried redundantly in the signal

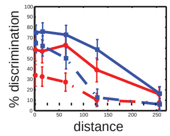
* features yielding the highest discrimination are different at short and long distances

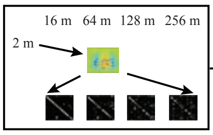


Templates

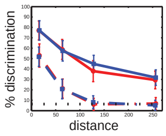


Very good results except at 256 m





Less interesting results at medium and long distance



Acoustic communication and sound degradation:
How does the individual signature of zebra finch calls transmit over distance?

Solveig C. Mouterde, Frédéric E. Theunissen, Julie E. Elie, Clémentine Vignal
and Nicolas Mathevon

Abstract

Background: Assessing the active space of the various types of information encoded by songbirds' vocalizations is important to address questions related to species ecology (e.g. spacing of individuals), as well as social behavior (e.g. territorial and/or mating strategies). Up to now, most of the previous studies have investigated the degradation of species-specific related information (species identity), and there is a gap of knowledge of how finer-grained information (e.g. individual identity) can transmit through the environment. Here we studied how the individual signature coded in zebra finch long distance contact call degrades with propagation.

Methodology: We performed sound transmission experiments of zebra finches' distance calls at various propagation distances. The propagated calls were analyzed using discriminant function analyses on a set of analytical parameters describing separately the spectral and temporal envelopes, as well as on a complete spectrographic representation of the signals.

Results/Conclusion: We found that individual signature is remarkably resistant to propagation as caller identity can be recovered even at distances greater than a hundred meters. Male calls show stronger discriminability at long distances than female calls, and this difference can be explained by the more pronounced frequency modulation found in their calls. In both sexes, individual information is carried redundantly using multiple acoustical features. Interestingly, features providing the highest discrimination at short distances are not the same ones that provide the highest discrimination at long distances.

Introduction

Birds' acoustic signals transmitted over large distances degrade in amplitude and in spectral and temporal structure as they propagate through the environment [1], [2]. These propagation-induced degradations reduce the active space of the signal, i.e. the distance from the emitter over which the information can be decoded by a receiver [3–6]. One of the challenges in songbird vocal communication is to investigate this acoustic active space, and more specifically to understand how degradation affects the message emitted by the sender as well as the response of the receiver(s).

Among various pieces of information that birds' vocalizations can potentially encode, cues about individual identity (individual signature) are particularly important. Indeed, individual recognition plays a fundamental role in male/female communication in the contexts of courtship behaviors and pair bond maintenance (especially in monogamous species) [7–9], as well as for territorial birds that need to recognize the vocalizations of strangers from neighbors in order to react accordingly [10], [11]. Moreover, individual recognition between parent and offspring is critical for breeding success, especially for colonial species that do not use fixed nest sites [12], [13]. The preservation of individual signature in propagated sounds raises an interesting question as it appears that individuality requires fine temporal and spectral information that may be highly susceptible to propagation-induced degradation [14].

To our knowledge however, only two studies providing an acoustic analysis of the long-range degradation of the individual signature in songbirds have been published. In the white-browed warbler *Basileuterus leucoblepharus*, the territorial song of the male consists of a succession of pure tones slowly decreasing in frequency. The emitter's individual identity is encoded in the first half of the song, which is high-pitched. After long range transmission, this first half, and consequently the individual signature, disappears due to the great susceptibility of high frequencies to degradation through the forest environment [14], [15]. Conversely, in the territorial call of the male corncrake *Crex crex*, the acoustic feature that is the most characteristic of each individual is the inter-pulse duration. This temporal code remains practically constant after propagating 100 m through dense vegetation, making it a good candidate for encoding the vocal signature at a distance [16].

In these two previous studies, the individual identity is coded either by frequency modulation of pure tones (white-browed warbler) or inter-pulse duration (corncrake), which makes the assessment of propagation-induced modifications quite simple. In the case of the white-browed warbler, the frequency-dependent attenuation over distance is responsible for the disappearance of the high-pitched notes, and for the corncrake the filling of inter-pulse silences by echoes is the main factor explaining the progressive loss of individual information. Conversely, the quantification of propagation-induced information loss of complex sounds displaying large frequency bandwidths together with amplitude and frequency modulations could be more problematic. In the present study, we focused on the distance calls of the zebra finch *Taeniopygia guttata*, which fit these characteristics.

Zebra finches are small gregarious songbirds from subarid regions of Australia that form strong pair bonds for life and live in large flocks in open grassy country with a scattering of trees and bushes [17], [18]. They use different types of vocalizations, one of the most frequently heard being the distance call. This loud call is emitted by both sexes and is used when birds establish acoustic contact at a distance [18]. It is a complex sound, consisting of a harmonic series modulated in frequency as well as amplitude. Distance calls differ between males and females [19], [20], the males' fundamental frequency being higher than the females' (typically 650-1000 Hz versus 500-600 Hz) as well as usually being shorter and more frequency-modulated (Fig. 1). It has been shown that the distance call bears an individual signature [21], and that zebra finches are capable of call-based individual recognition [8], [9], [19], [22], [23]. The "active space" of these calls has been estimated based on naturalistic observations to be up to 100 meters [18]; similar conclusions were reached with theoretical calculations using discrimination thresholds for masked non-degraded signals in this species [4]. However, no experimental work has yet investigated the long-range transmission of the individual signature in the distance call.

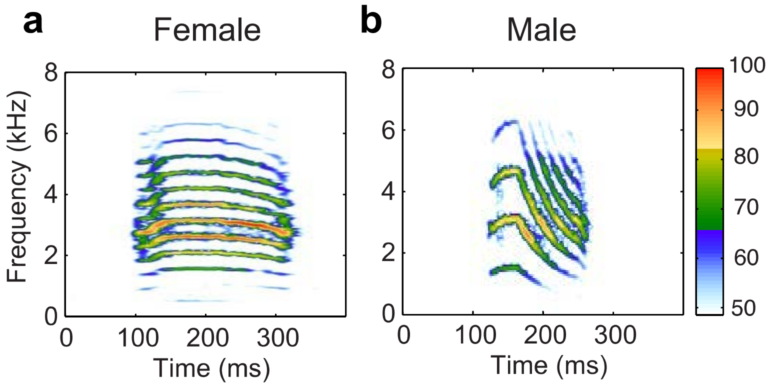


Figure 1. Spectrograms of a female (a) and male (b) distance call, recorded at 2 m. Both signals were high-pass filtered above 500 Hz to avoid displaying low-frequency background noise. The color scale is in relative dB as shown on the color bar with 100 dB corresponding to maximum amplitude observed.

Here we characterized and quantified the degradation of female and male zebra finch calls during propagation, focusing on the acoustical features that carry information about individual identity. We performed this descriptive analysis to answer three sets of questions. First, what are the acoustical features that carry the most information about the individual signature, and are these most informative features the same at different distances? Second, how would a discrimination performance be affected if the animal based its discrimination on a single and distance independent set of acoustical features instead of using different sets of features adapted to each tested distance? Finally, are there significant differences in the individual signature of female vs. male calls in the way they degrade with propagation?

Materials and Methods

We performed two sets of propagation measurements: the first set was performed in Australia in a typical environment of the wild zebra finch, and the second set was performed in France in an open field. The measurements in Australia were performed with a limited database of zebra finch calls and can be considered as a pilot study for characterizing the acoustic properties of the species' natural environment. The background noise distributions and the frequency transfer functions obtained in both environments were compared to assess both the ecological relevance and some of the limitations of these analyses. The propagation experiments performed in France included multiple call exemplars of each individual, and aimed to fully characterize the acoustic features that carried information about the caller's identity.

Ethics statement

The experimental protocol for the recording of distance calls of our zebra finch subjects was approved by the Jean Monnet University's animal care committee (authorization n°42-218-0901-38 SV 09 to the ENES lab).

Calls database used in the experiments and recording locations

For the pilot experiments performed in Australia, we used calls from both wild and domesticated zebra finches (16 calls used: 2 samples each from 2 wild males, 2 wild females, 2 domesticated males and 2 domesticated females). The domesticated zebra finches had been previously recorded in France (see protocol below). The calls from the wild birds were recorded at the University of New South Wales Arid Zone Research Station at Fowlers Gap, 112 km north of Broken Hill in far western New South Wales, Australia (Gap Hills; 30°57'S, 141°46'E). Here, a large zebra finch colony is established, with birds breeding both in natural vegetation and in nest-boxes [24]. We recorded calls using omnidirectional tie-microphones (AKG C417) placed inside nest-boxes and connected to a portable solid state recorder (Marantz PMD670; sampling frequency: 44100 Hz; D&M Holdings Inc., Kanagawa, Japan). These recordings were conducted under the authorities of the Animal Ethics Committees at the University of New South Wales and Macquarie University and a Scientific Research Permit from the New South Wales Parks and Wildlife Service.

For the propagation measurements performed in France, we used calls from 32 domesticated zebra finches (16 distance calls from each of 16 females and 16 males = 512 calls). The domesticated birds were bred in the ENES laboratory aviary (14L/10D photoperiod with adapted wavelengths; food and water *ad libitum*; temperature between 23 and 25°C). They were recorded in a soundproof room using a Sennheiser MD-42 microphone placed 0.2 m above the cage and connected to a Marantz PMD 670 recorder (sampling frequency: 44100 Hz). Each bird was recorded in the presence of two female zebra finches placed 3 m away and used as an audience to minimize stress [25]; the bird was stimulated with playbacks of distance calls that had previously been recorded from conspecific birds. During the recording, conditions of temperature, food and water availability were the same as in the aviary. All the calls were normalized by matching the maximum values of the sound pressure waveforms using Goldwave (version 5.57), prior to the propagation experiments.

Recording and processing of propagated calls

1. Sound transmission experiments in Australia

The transmission experiments in Australia were performed in November 2008 at two locations, next to the research station. The first experimental location was an open environment deprived of vegetation. On this site, experiments were performed during the day with light wind. The weather was clear and the temperature was around 30°C (Australia site 1). The second location was a subarid environment with sparse vegetation including small trees, the landscape being open with little relief [18]. Experiments at this location were performed on a completely windless evening. The weather was clear and the temperature was around 25°C (Australia site 2).

For both Australian recordings, the sounds were played back via an amplified loudspeaker (Minivox Lite PA System, Anchor Audio, Inc.) connected directly to a laptop computer. The loudspeaker was placed at a height of between 1 and 1.3 m. The volume of the speaker was set such as to match the intensity typical of zebra finch calls, around 70 dB at 1m [9]. Sounds were recorded using an omni-directional microphone (Sennheiser MD 42) connected to a Marantz PMD670 recorder (sampling frequency: 44100 Hz). The microphone was held at 1 m above the ground. We recorded the calls sequences 2 m, 5 m, 10 m, 20 m, 50 m and 100 m away from the speaker.

2. Sound transmission experiments in France

The transmission experiments in France were performed in October 2010 in the afternoon on a level field in Bellegarde-en-Forez, Loire, France. This recording site was chosen to match the open environment of the Australian desert. The weather was cloudy, there was little wind and the temperature was 11°C. The complete calls database was broadcast from a portable solid state recorder (Marantz PMD671) connected to a MegaVox speaker (PB-35W) placed on a stool, 1.3 m high. The speaker volume was set to obtain a sound level of 70 dB SPL at 1 m (Velleman Sound Level Meter DVM 1326). The sounds were recorded with a Schoeps microphone (MK4 cardioid, on a CMC6-U base) equipped with a Schoeps Basket-type Windscreen (W 20) and set 1.30 m high. The microphone was connected to a second Marantz recorder (PMD671; sampling frequency: 44100 Hz). We recorded the calls 2 m, 16 m, 64 m, 128 m and 256 m away from the source, twice for each distance.

3. Processing of propagated signals

We used a custom-made Matlab (Mathworks) script to cross-correlate the propagated recordings with the original sequences. For the recordings performed in France, we then compared the signals of the propagated calls from the two recording sessions, by ear or by a visual assessment of spectrograms when necessary, and selected the better of the two recordings for each call and at each propagation distance. This selection allowed us to exclude recordings that had been impaired by unexpected transient sounds that were not relevant to the study (e. g. birds calling in the vicinity).

Acoustic analysis

1. Visualization of the spectro-temporal properties of the environmental noise and the propagated calls

a. Frequency spectra. Using custom-made Matlab scripts, we calculated the frequency spectra (FS, in dB units) for all calls from the same sex at each propagation distance. We also calculated the FS of the background environmental noise that was present during these recordings.

b. Spectral transfer function. Spectral transfer function can be used to quantify signal attenuation (or gain) and phase shifts as a function of frequency for a given propagation distance. Here we calculated, the gain of spectral transfer function by normalizing the cross-spectrum between a propagated sound and the reference sound by the FS of the reference sound. We used the recordings obtained at 2 m as the reference sounds, and the recordings at 50 m (Australia) and 64 m (France) as the second propagated sound. In other words, we estimated the gain of the transfer function resulting from 48 m of propagation in Australia and 62 m in France. Estimates of the standard error in the gain function were obtained using the Jackknife resampling technique [26].

c. Modulation power spectra. In order to visualize the joint spectro-temporal modulations of the distance calls propagated in France and their evolution with propagation, we calculated the modulation power spectra (MPS) of the calls: the joint second order statistics of the spectro-temporal amplitude envelopes obtained from a spectrographic representation of the sound [27]. The MPS is obtained as follows. We first calculated the spectrogram of each call using a Gaussian window (symmetric in time and frequency domains) 70 Hz wide in the frequency domain or 2.27 ms wide in the time domain. The MPS is then simply the 2-D power spectrum of the log spectrogram [27]. The time-frequency scale of the spectrogram (70Hz) determines the spectral and temporal Nyquist limits of the modulation spectrum (7.14 cycles/kHz for spectral modulations and 220 Hz for temporal modulations). MPS were obtained for each propagation distance and each sex.

2. Parameters used for the statistical analysis of individual signature

We used two distinct sets of acoustic parameters to test for the presence of a vocal signature in the propagated calls: (a) a set of parameters that separately describe the amplitude in the spectral domain (the spectral envelope) and the amplitude in the time domain (the temporal envelope, often simply called amplitude envelope); (b) the spectrogram. Since our spectrographic representations are invertible, using a complete spectrogram not only circumvents the use of *a priori* assumptions on the nature of the information-bearing acoustical features but also provides an upper bound for discriminability.

Prior to these analyses, the sounds were band-pass filtered between 0.5 to 8 kHz in order to reduce irrelevant environmental background noise. Those frequency cutoffs were chosen based on the zebra finch's audiogram [28].

a. Parameters used to describe the separate spectral and temporal features. We extracted the spectral amplitude envelope (amplitude as a function of frequency) and temporal amplitude envelope (amplitude as a function of time) of each call. Each amplitude envelope (spectral and temporal) was then converted to a density function by dividing each value of amplitude by the sum of all amplitude values. We quantified the shape of these normalized envelopes by estimating the moments of the corresponding density functions: their mean (i.e. the spectral centroid for the spectral envelope and temporal centroid for the temporal envelope), standard deviation (i.e. spectral bandwidth and temporal duration), skewness (i.e. measure of the asymmetry in the shape of the amplitude envelopes), kurtosis (i.e. the peakedness in the shape of the envelope) and entropy. The entropy captures the overall variability in the envelope; for a given standard deviation, higher entropy values are obtained for more uniform amplitude envelopes (e.g. noise-like broad band sound and steady temporal envelopes) and lower entropy values for amplitude envelopes with high amplitudes concentrated at fewer spectral or temporal points (e.g. harmonic stacks or temporal envelope with very fast attack and decay). The spectral envelope was obtained with the Welch's averaged, modified periodogram estimation of the power spectral density using a Hann window of 23 ms and an overlap of 99%. The temporal envelope was obtained by rectifying the sound pressure waveform and low-pass filtering below 50 Hz. With these procedures, we obtained 10 acoustical parameters, 5 describing spectral features and 5 describing temporal features. Since these parameters had different units, Z-scores were calculated prior to using them in the multivariate discriminant analyses.

b. Full spectrographic representation. As stated above, we calculated an invertible spectrogram of each call using a Gaussian window and a time-frequency scale of 70 Hz-2.27 ms. Because the dimensionality of this representation was higher than the total number of calls in our database, we used a Principal Component Analysis (PCA; using the princomp function of Matlab) for dimensionality reduction. The discriminant analysis was then performed using the coefficients of a subset of the principal components.

These two complementary approaches enabled us to compare the acoustical nature of vocal signatures, using discrete and more easily interpretable envelope parameters that have proven to be useful, if somewhat subjective, in the investigation of information-bearing features [29], as well as more complex parameters (principal components coefficients) extracted from a complete and invertible representation of the signals. To be concise, we will refer henceforth to the 5 pairs of separate spectral and temporal parameters as the “envelope parameters” and to parameters describing the spectrogram as the “spectrogram principal component parameters”, or more simply “SPC parameters”. The extraction method for each set of parameters is summarized in Fig. 2a.

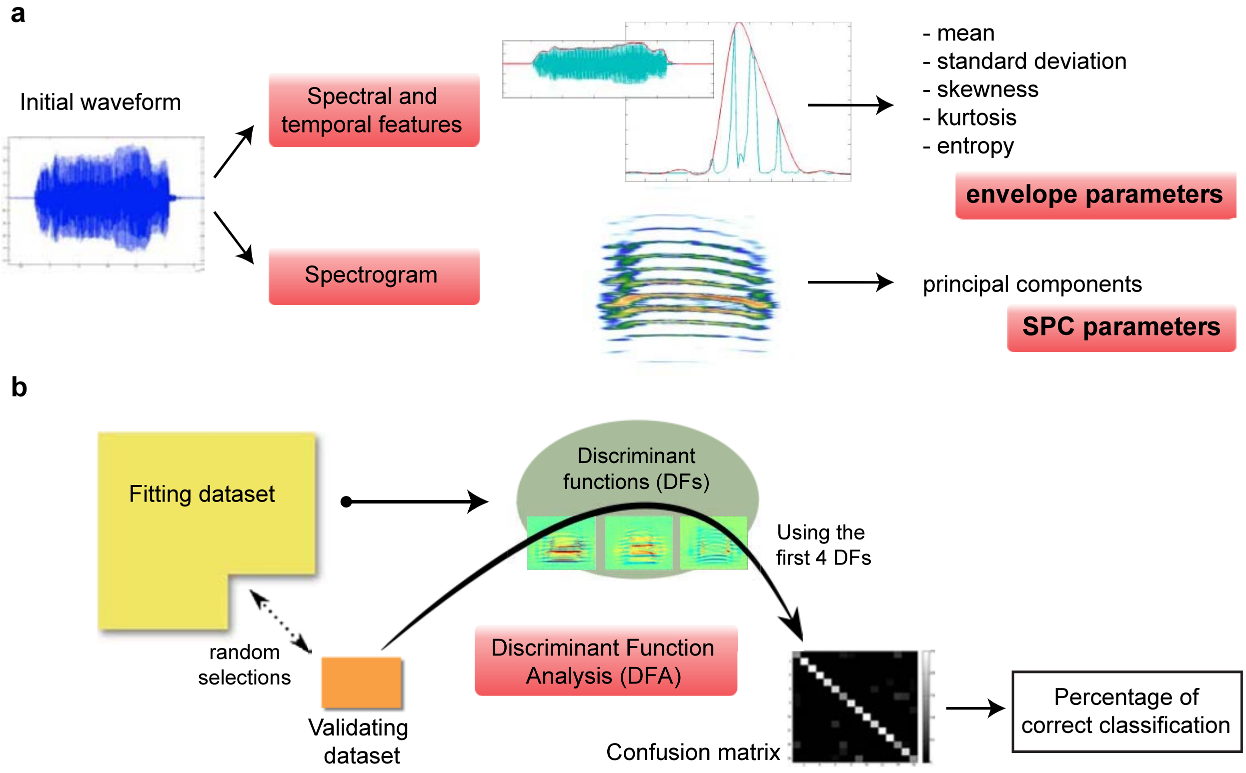


Figure 2. Overview of the acoustical analysis. This figure summarizes the procedure for extracting both sets of parameters (a), and the subsequent discriminant function analysis (b). The discriminant functions were calculated using the fitting dataset, and we tested the validating dataset against these to obtain the percentages of correct classification.

3. Statistical analysis of the vocal signature in calls

To test for the presence of an acoustic signature in the propagated distance calls, we used cross-validated linear discriminant function analyses (DFA), performed separately for each sex. The DFA finds the linear combinations of parameters that maximally separate the data points (the calls) that belong to different categories (the birds identities); these linear combinations are called discriminant functions (DFs) and the significant DFs define a Euclidian subspace where calls can be projected. The cross-validation consists of first calculating a set of discriminant functions using a fitting dataset, and then testing these functions on a validating dataset (non-overlapping with the fitting dataset) by assessing the number of calls that had been correctly classified for each individual, the percentage of correct classification (Fig. 2b). We performed this DFA using either the 10 envelope parameters or the SPC parameters.

Prior to the final calculations, we performed a series of preliminary analyses to determine the number of principal components (PCs) used for the SPC parameters and the number of discriminant functions (DFs) used for quantifying the discrimination for both the envelope and SPC parameters. We found optimal values for these hyper-parameters using the data at 2 m as a reference, and applied them for all other conditions. We first aimed to define the number of PCs to be used to reduce the dimensionality of the data while representing a reasonable amount of variance; to do this we estimated the percentages of correct classification using the 10 first significant DFs (at 2 m using the SPC parameters, 13 were significant for the females, and 15 were significant for the males as assessed by the Wilks Lambda test statistic) and a varying number of PCs from 10 to 100. We found that using 25 PCs captured most of the information about identity as witnessed by a saturation of the percentages of correct classification around that number (94.9% with 25 PCs and 95.4% with 100 PCs, using 10 DFs). We then compared the eigenvalues for each DF (fixing the number of PCs to 25) and found that using the first 4 DFs (DF1 to DF4) was an appropriate tradeoff, as it accounted for 83.9% for the females and 82.6% for the males of the variance between birds. Furthermore, in all our discriminant analyses (described below), we found at least 4 significant DFs. We were thus able to always use the same number of DFs in all our calculations and directly compare results. Finally, we assessed the number of cross-validation iterations required to obtain robust values for the percentage of correct classification; we found that 250 iterations were largely sufficient to reach stable values.

We quantified the discriminability by calculating the percentage of correct classification. We also determined which acoustical features were the most important for individual discrimination by examining the form of the significant DFs (for the SPC parameters) or the effect of each set of parameters independently (for the envelope parameters). The discriminant analyses were only performed for the more complete data set of propagated calls obtained in the experiments in France.

4. Extraction of the most important parameters for individual discrimination

A different procedure was used for identifying the relative importance of the envelope parameters and the SPC parameters in the individual discrimination task. For the envelope parameters, we simply repeated the DFA with only the temporal envelope parameters or only the spectral envelope parameters. For the SPC parameters, we represented the first 4 DFs in the spectrographic space: since both the discriminant analysis and the PCA decomposition of the spectrogram are linear operations, the inverse rotation and scaling can be applied on the discriminant functions to represent them in a spectrographic space, allowing us to describe them as we would on a spectrogram.

5. Comparisons between the percentages of correct classification

To compare the results obtained with different set of parameters, at different distances and for both sexes, we performed analyses of covariance (ANOCOVA), using a general linear model framework (Statistical Toolbox in Matlab), with the percentage of correct classification as the dependent variable, the sex (male vs. female) and the type of parameters used (SPC vs. envelope) as factors, and the propagation distance as a covariate.

Results

In a first step, we will describe the frequency dependent spectral changes in Australia and in France. In a second step, we will examine the changes in the temporal-spectral modulations due to propagation. In the remaining sections, we will then examine the structure in the calls that carries the information about identity, how this information changes for propagated calls, and whether these changes can be explained by the degradations.

Spectral degradation in propagated signals in France and Australia

On Fig. 3, we illustrate the effect of propagation on signal quality by showing the spectrograms of the same female (a) and male (b) calls recorded at 16 m, 64 m, 128 m and 256 m in the open field in France.

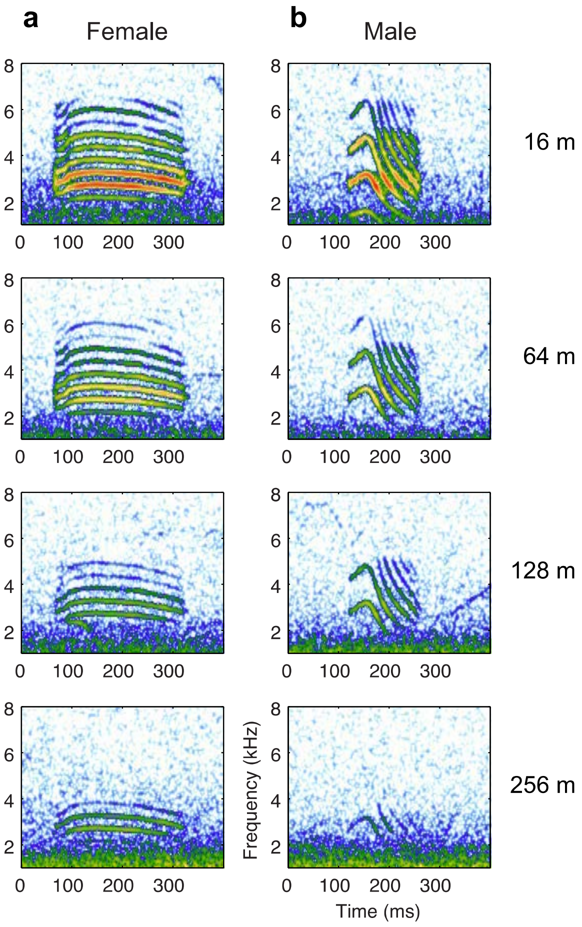


Figure 3. Spectrograms of the same female (a) and male (b) distance call recorded at various distances. All signals were high-pass filtered over 500 Hz, and the scale was applied to all spectrograms.

The propagation-induced band-pass filtering of the signal is clearly shown, as well as the significant reduction of the signal-to-noise ratio (SNR) with distance. Specific higher-order spectral structures such as pitch for the female call become also less salient as the fundamental drops below the noise level at low frequencies and fewer harmonic components are available at higher frequencies. Similarly higher order temporal structures, such as the relative durations of the upsweep and downsweep in the male call, are also affected. Fig. 4 shows the mean frequency spectra (FS) of propagated calls for both females and males. The curves quantify the attenuation of the signal with increasing distance. The band-pass profile of the calls is clearly visible at 2 m, with a large peak centered on about 3 kHz for both sexes. However and while the signal is clearly above the noise floor up to 8 kHz at 2 m, this upper frequency bound is reduced to 6 kHz at 64 m and around 4.4 kHz at 128 m. Similarly, the lower frequency bound increases as a function of propagation distance. At further distances, signal energy is only found in a narrow peak centered at 3 kHz.

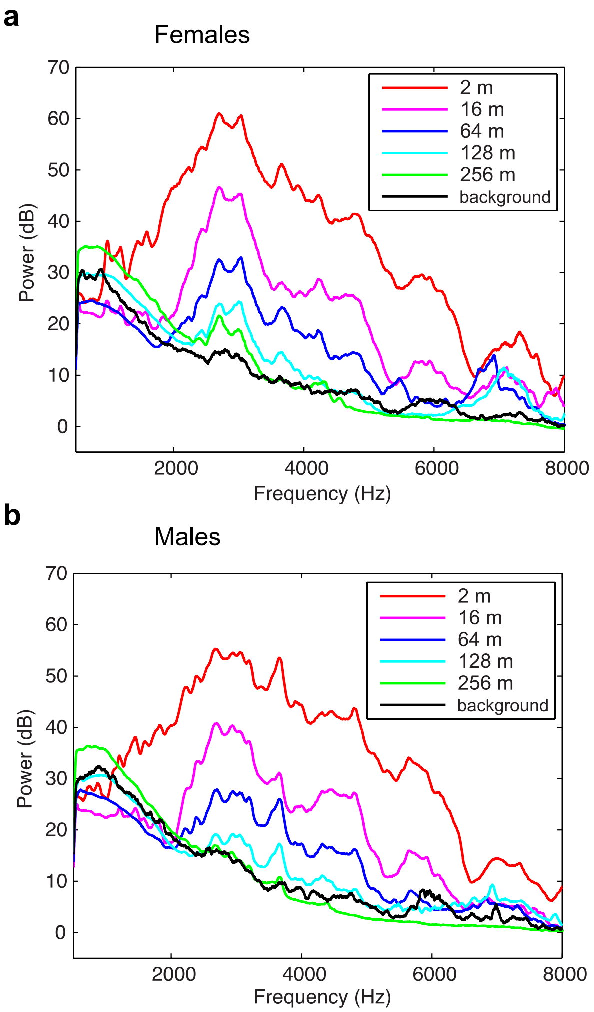


Figure 4. Frequency power spectra of calls at every propagation distance (from 2 m to 256 m). These were calculated using all female (a) and male (b) distance calls recorded in France. The frequency spectrum of the background noise at the recording sites is shown in black. This curve was obtained by averaging the noise spectrum across all recording sites and can therefore be slightly higher than the noise floor at particular distances, as it can be observed for the lower frequency range.

To what extent are these measurements affected by the environment? On Fig. 5, we show the frequency transfer function between 2 and 64 m for calls recorded in France and between 2 and 50 m for the sounds recorded in Australia. The transfer function between a call recorded at 2 m and the same call recorded at 64 m shows the attenuation of the signal as a function of frequency, taking into account all environmental effects, both landscape and noise conditions. The frequency transfer function in France is relatively flat between 1 and 7 kHz, at approximately 10 dB below the 30 dB of attenuation expected from spherical spreading (horizontal line on Fig. 5). Approximately 5 dB can be attributed to air damping (ISO 9613-1:1993) and thus, only 5 dB can be attributed to other factors such as wind, terrain and vegetation. The transfer functions obtained from the recordings performed at site 2 in Australia have more structure, with variations alternating between 5 and 20 dB below the value expected from spherical spreading. We attribute most of the variations to the sparse vegetation that was found at that location. The recordings at site 1 were performed in a more open region and the corresponding transfer function shows different fluctuations at low frequencies and a greater attenuation at high frequencies.

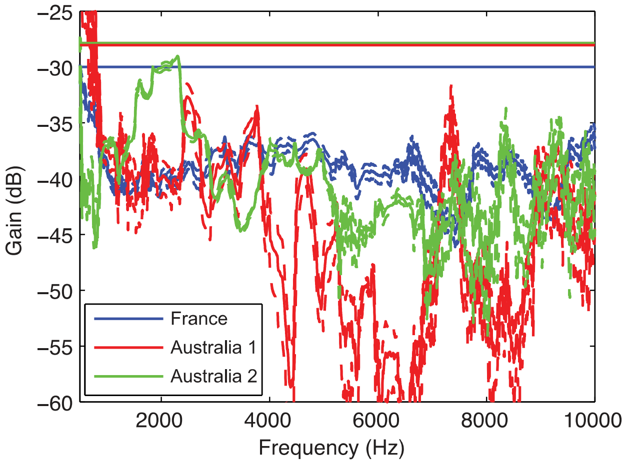


Figure 5. Comparison of transfer functions in France and in Australia. The transfer functions were calculated between the recordings obtained at 2 m and those recorded at 50 m in Australia and between 2 m and 64 m for the recordings made in France. The horizontal lines represent the theoretical attenuation due to spherical spreading only for each site. The Australia site 1 was an open area close to the nesting sites and the recordings were performed during the day with low wind. The Australia site 2 is next to zebra finch nesting sites and includes sparse vegetation and the recordings were performed on a windless evening. The recordings in France are in an open field and were obtained during the day in low wind conditions (see text for more details).

These measurements show, as it is well known, that the attenuation of sound in a natural environment will depend on the landscape. However, the deviations from spreading and atmospheric absorption observed in Australia are relatively small and, because of the sparse vegetation, idiosyncratic to very local conditions. For these reasons, the frequency transfer function of the terrain in France can be considered to be a relatively good (vegetation independent) average of what we have measured in Australia. The rest of the results will focus on the analyses performed with the recordings obtained in France.

The modulation power spectrum of degraded calls

While the spectrograms and frequency power spectra of the propagated calls illustrate the changes in SNR as function of frequency and distance, they fail to quantify how this attenuation affects potentially informative spectral and temporal structures in the signal. To quantify these structural changes, we calculated the modulation power spectrum (MPS) of the calls, which is the 2-D power spectrum of the joint spectral-temporal envelope of the sound [27] (see Methods). On Fig. 6 we show the average MPS of all calls at each distance, for each sex. Temporal modulations are shown on the x-axis and characterize the power in the temporal amplitude envelope as a function of temporal frequencies (in Hz); signals that contain fast changes in their envelope (such as a quick onset or a fast chirp) are characterized by power at the higher temporal modulation frequencies. The spectral modulations are shown on the y-axis and characterize the power in the spectral envelope in terms of spectral frequencies (in cycles per kHz); signals with clearly defined harmonics show power at the spectral modulations corresponding to the inverse of the pitch frequency (higher for lower pitch). Noisy broadband sounds show broad power at low spectral modulations. Any point away from these axes represents joint spectro-temporal modulations such as upsweeps (in the area with negative temporal modulation) and downsweeps (in the area with positive temporal modulation) [27].

For the females (Fig. 6a), the MPS at short distances (2 m and 16 m) show a large area of energy around 2 cyc/kHz and very low temporal modulations (centered at 0 Hz), corresponding to the salient and relatively steady pitch of the call; the spectral modulation value of 2 cyc/kHz corresponds to a fundamental frequency of around 500 Hz, consistent with the mean fundamental frequency of the female distance calls used in our dataset. The spread of energy at 2 cyc/kHz along the horizontal dimension represents the up- and down-sweeps at the beginning and end of the calls. The smaller peak of energy around 1 cyc/kHz is characteristic of differential energy in the odd vs. even harmonics (1 cyc/kHz corresponds to a frequency of 1 kHz, which characterizes a boost in energy every second harmonic) and energy below 0.5 cyc/kHz corresponds to the spectral envelope produced by the filtering of the upper vocal tract (the equivalent of the formants; Fig. 6c).

Fig. 6a clearly shows that for the females, the joint spectro-temporal modulations corresponding to the up and down sweeps are progressively lost with propagation distance, while the energy in the pitch area remains visible at 256 m, albeit less prominently. These degradations are also evident on the spectrogram for the example call in Fig. 3: the faster modulation changes found at the beginning and end of the female call fall more rapidly below the noise ceiling than the louder and less dynamic central section of the call. The relative decrease of energy in the pitch area is a consequence of the loss of the lower and higher harmonics as well as a decrease in the sharpness of the remaining harmonics in the 2-4 kHz region.

For the males (Fig. 6b), information about pitch is also clearly visible, between 1 and 2 cyc/kHz, which is consistent with the higher fundamental frequency of their calls around 700 Hz. Similarly to what is also observed in females but to a much greater extent, the energy corresponding to the up- and down-sweeps is visible. In particular, a large bulge of energy is found on the right side of the y-axis corresponding to the fast down-sweeps that characterize the trailing end of the male call. While the information about up-sweeps is progressively lost (as was the case for the females), these high frequency down-sweeps, with a modulation rate of about 25 to 40 Hz, remain clearly visible up to 256 m. These modulations, characteristic of the male call, preserve the sexual dimorphism of the calls at all distances. Finally, for both males and females, the energy visible along the x-axis (corresponding to the overall temporal envelope) is progressively reduced to lower modulation frequencies with distance; thus, sound propagation also degrades the faster dynamics in the temporal envelope.

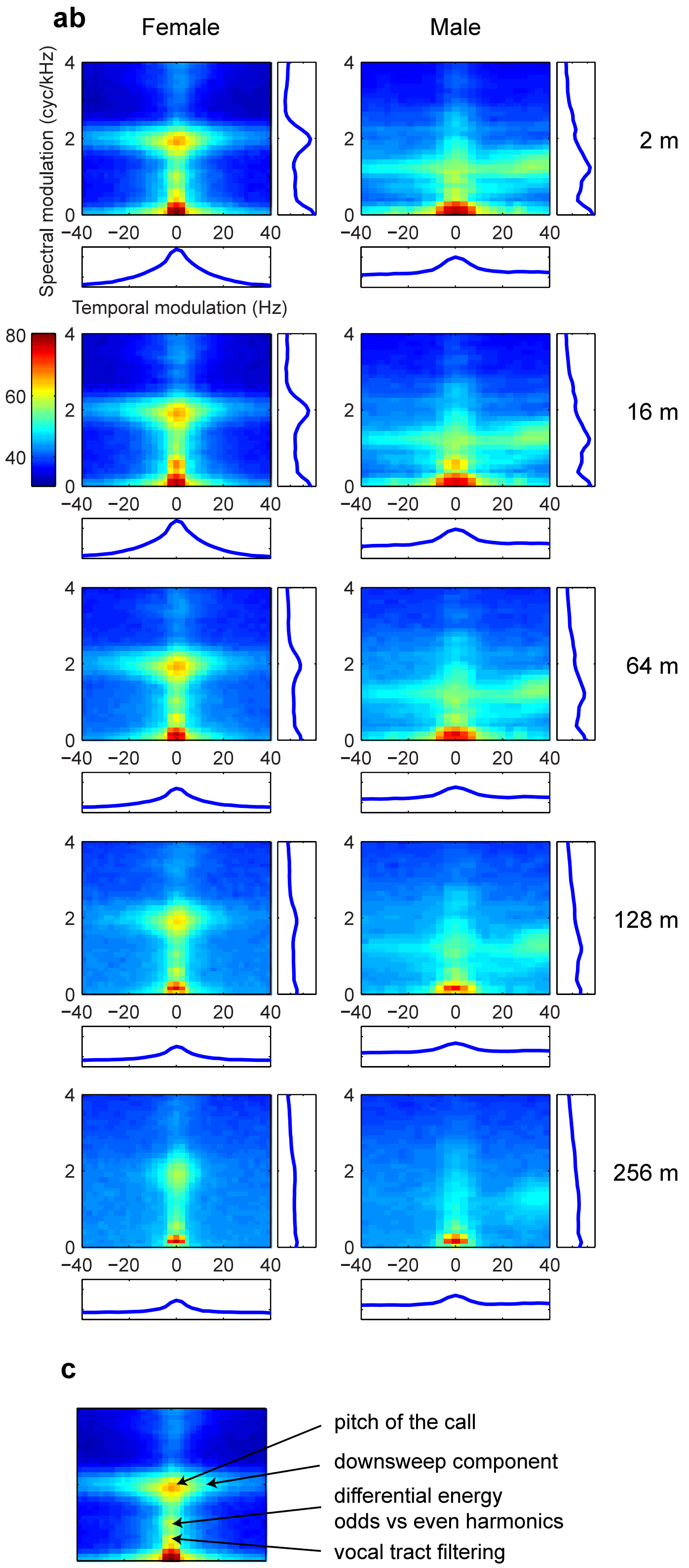


Figure 6. Modulation power spectra (MPS) of female (a) and male (b) calls for each propagation distance. Mean MPSs are shown for each propagation distance used in this study (from 2 m to 256 m). Projections on the spectral modulation axis and the temporal modulation axis are shown beside and below each MPS respectively. The major observable features are shown and labeled on one example shown in panel (c).

Effect of propagation on individual vocal signatures

The nature of the individual information in the distance call and the way this information degrades with distance were assessed and quantified using a cross-validated linear discriminant function analysis (see Methods and Fig. 2). These discriminant analyses were performed under three scenarios. First, we analyzed the data separately for each distance, in order to obtain ceiling values for discrimination through quantifying correct classification in validating datasets also obtained from the corresponding distance. This analysis simulated a scenario where the bird could use the actual distance of the sound source as independent information to discriminate individuals. We then conducted two additional analyses that allowed us to determine putative discrimination abilities in simulated scenarios where the bird would not have access to independent information on the actual distance of the sound source. In one scenario, we tested a model where the bird uses the acoustical features most useful for identity discrimination at short distance for all of its discriminations: we used the calls recorded at 2 m as a reference for analyzing calls recorded at all other distances. In the other scenario, we tested a model where the bird also uses the same acoustical features at all distances but where these features are found by taking all propagation distances into account: we used all calls recorded irrespective of distance (and ignoring this information) for analyzing individual discrimination.

1. Discriminant analysis for each distance separately

On Fig. 7a, we show the percentages of correct classification as a function of distance, for both sexes and both sets of parameters (envelope and SPC) when the discriminant analysis is performed for each distance separately. Using the envelope parameters, the percentages of correct classification at 2 m average 80.2% for male calls (n = 16 individuals), and a noticeably lower value of 62.6% for female calls (n = 16 individuals). These values increase when the SPC representation is used, averaging 82.6% for the males and 76.5% for the females (Fig. 7a – values at 2m). These classification values remain above chance for most other distances. However, when the envelopes parameters are used to describe the calls, the percentage of correct classification obtained at 256 m is close to chance level for both male and female calls. Statistical analysis shows that there is a significant effect of distance (ANCOVA with the percentage of correct classification as the dependent variable, sex and type of parameter as factors and distance as a covariate; $p < 10^{-6}$). Discrimination is thus more and more difficult with increasing distances (with an average decrease of 1.94 % points every

10 m). There is also an effect of sex on the percentage of correct classification, the discrimination of female calls being on average 9 percentage points below those of males ($p = 0.02$). Finally, there is a significant effect of the type of parameter, with the envelope parameters yielding worse performance by 21 percentage points ($p < 10^{-4}$). When interactions are included in this linear model, one also finds that the percentage of correct classification decreases more rapidly with distance for the envelope parameters than for the SPC parameters ($p = 0.009$). Further analysis taking into account each sex separately shows that this difference between the two slopes is due to the males only (F = 10.24, $p = 0.0186$): the performance of the discrimination deteriorates at a faster rate for the first distances (up to 128 m) with the envelopes parameters compared to SPC parameters.

In summary, this analysis shows that the males' individual signature is stronger than the females' and that for both sexes, using the SPC parameters give significantly better results in terms of discriminability than using the envelope parameters, especially at longer propagation distances.

2. Discriminant analysis taking 2 m as a reference: maximizing peak discrimination

Here, discriminability between individuals decreases more rapidly with distance than for the previous analysis, the slope of the curves being less linear (Fig. 7b). Specifically, discriminability decreases on a steeper curve at medium to long range for the SPC parameters but still remains well above chance level at 256 m for both sexes. For the envelope parameters, the percentages of correct classification at short range are well below those obtained in the previous analysis, and by 128 m discriminability is equivalent to chance for both sexes. Indeed, the ANCOVA analysis using the percentage of correct classification as the dependent measure, the type of parameter used in the DFA and the sex as factors and distance as the covariate showed a significant main effect for the type of parameters ($p < 10^{-4}$) with an estimate decrease of 30 percentage points for the envelope parameters relative to the SPC parameters, but no significant interaction between the type of parameter and sex. In this case, we failed to find a significant main effect for sex.

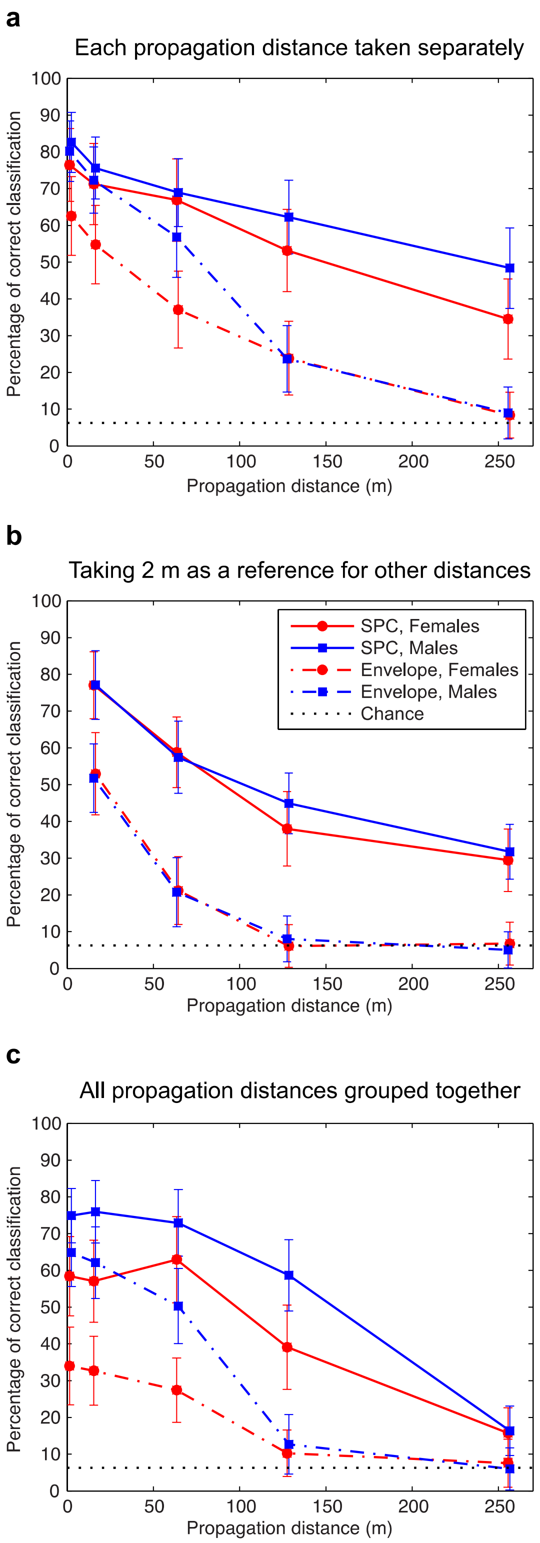


Figure 7. Mean percentages of correct classification obtained for each propagation distance and both sexes. (a) Taking the data for each propagation distance separately, (b) taking the 2 m data as a reference for other distances, and (c) taking all propagation distances into account. SPC parameters are 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.

3. Discriminant analysis of all propagation distances grouped together: maximizing average discrimination

The percentages of correct classification shown on Fig. 7c quantify the discriminability obtained using a single set of DFs based on data from all distances. As it is the case when the DFA is performed for each propagation distance separately, the ANCOVA shows that there is a main effect of distance, sex and parameter type. Moreover the estimated values are very similar (compare 7c with 7a): the increase in distance results in 1.9 percentage points loss per 10 m ($p < 10^{-6}$), discrimination of male calls is superior than female calls by 14.9 percentage points ($p = 0.002$) and envelope parameters give percentages of correct classification that are on average 22.4 points worse than the discrimination obtained with the SPC parameters ($p < 10^{-4}$). Thus, this strategy results in relatively small penalties relative to the strategy that involves using the optimal DFs at each distance, but only when the SPC parameters are used (compare with 7a): the difference in the percentages of correct classification was within our estimate of the standard error for all distances except 256 m. For the envelope parameters, the decrease in performance for the males call was also small and within the estimated standard error, while for the females we notice substantially lower discrimination rates at short distances (2 m and 16 m).

These results show that as long as the SPC parameters are used, a single distance-independent template can give very good results in terms of discriminability, with a noticeable drop occurring only at 256 m. Thus, although different acoustical features (as will be described in more detail below) could be used to optimize discrimination for different propagation distances, we also found a strategy using a single and robust set of acoustical features that yields good discrimination at all distances.

4. Acoustical parameters used for identity discrimination

As can be seen on Fig. 8, the spectral envelope carries more information than the temporal envelope for females at 2, 16 and 64 m while the opposite is true for males. At 128m, spectral and temporal envelope features are equally informative in females whereas spectral envelope features start carrying more information than temporal features for males. At 256 m, temporal features carry no more information for both sexes. In brief, informative features follow a clear sexually dimorphic pattern with not only differences in the relative importance of spectral vs. temporal features between the sexes but also opposite trends as a function of increasing propagation distance. This analysis also sheds light on the redundancy in the information in the temporal vs. spectral envelope. If the information was completely independent, the spectral and temporal bars shown on Fig. 8 would add up to 100%. The data show that there is always some redundancy in the spectral and temporal envelope information but 1) that this redundancy is higher in males that in females and 2) that it decreases for males while it increases slightly for females as propagation distance increases.

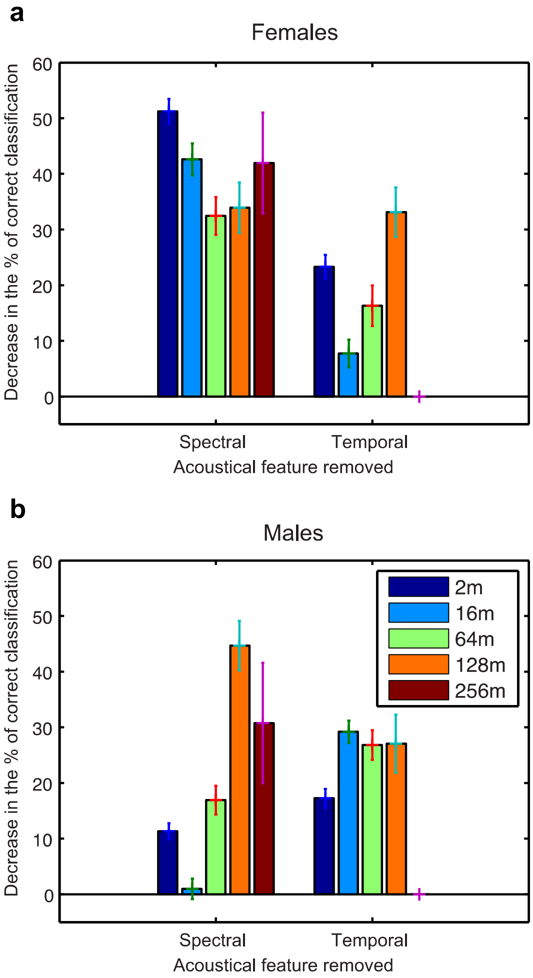


Figure 8. Comparison between the importance of spectral and temporal features of the envelope parameters. Calculations are shown separately for females (a) and males (b). Histograms represent the decrease in the percentage of correct classification observed when removing the spectral (left) or temporal (right) parameters from the DFA. Error bars show two-standard errors estimated by bootstrapping. For each distance, except for 128 m for females, the changes observed by removing spectral information vs. the temporal information are significantly different (all p values < 10^{-5} using a two-sided t-test).

A more explicit description of the effect of propagation on the informative features can be obtained with SPC parameters, by analyzing the evolution with distance of the DFs represented in the spectrographic space (Fig. 9). One striking result of that visualization is that both the range of the frequency spectrum and the duration in time where intensity is available shrinks as distance increases. While a large range of the frequency spectrum (between approximately 1 and 6 kHz) can be used to extract the individual signature at 2 m, the frequency band of the signal then shrinks as a function of distance and becomes restricted to a narrow range between 2 and 4 kHz at 256 m. The decrease in the temporal range is most evident for DF3 and more pronounced for the male calls: while the temporal range of the male DF3 for 64 m is above 200 ms it drops to below 100 ms at 256 m.

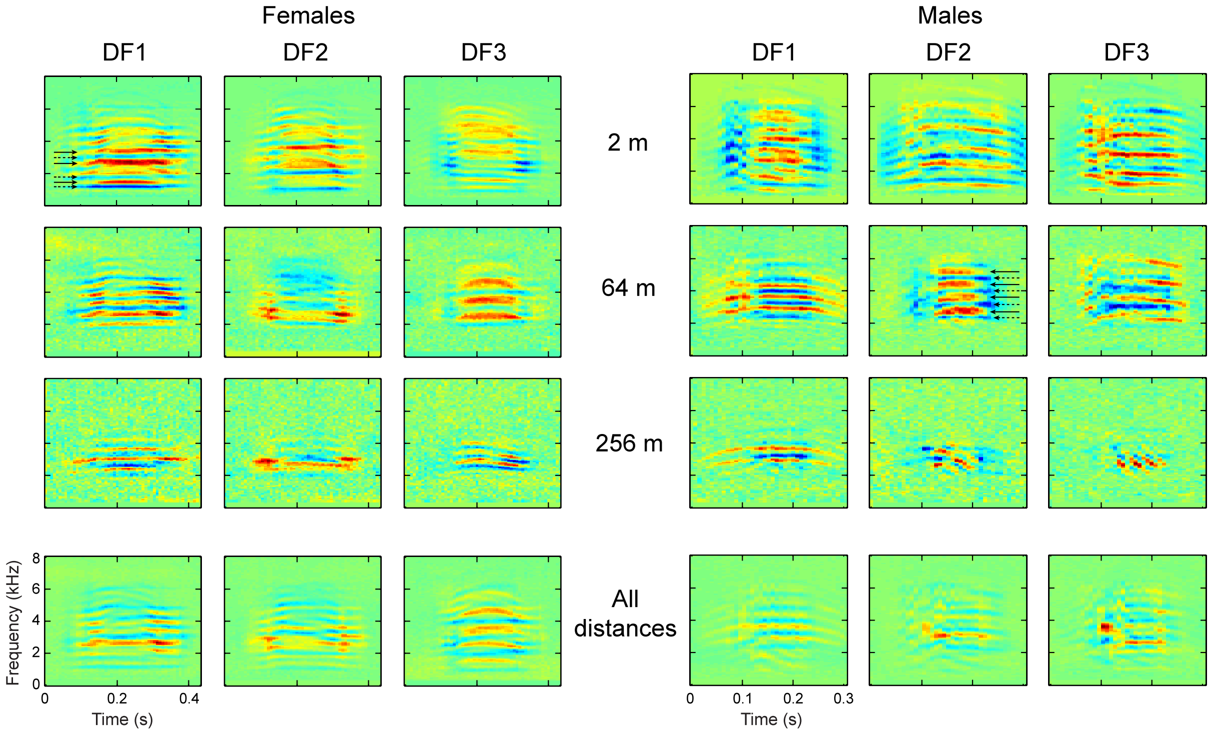


Figure 9. Representation of the discriminant functions (DFs) projected into the spectrographic space for the SPC parameters. The first 3 DFs, obtained from the SPC parameters, are shown for females (left) and males (right). Each row indicates the type of dataset used to perform the DFA: each distance taken separately for 2 m, 64 m and 256 m (first 3 rows) and all distances taken into account (bottom row). This representation enables a description of the most important features in the spectrogram that can be used to discriminate between individuals at various distances.

The DFs in Fig. 9 also show that fine spectral cues (the narrower alternating blue and red bands in the frequency bands) appear to be useful at all distances, and for both sexes. As explained above, this spectral structure is related to the pitch of the call. Thus, even though the bandwidth which can be used to extract this pitch information is reduced, this analysis shows that pitch information remains robust, a testament to its redundant character in broadband harmonic stacks. One feature that strikingly sets sexes apart at long range is the selectivity for downward sweeps that appears for males in DF2 and DF4 at 128 m (not shown on the figure) and in DF2 and DF3 at 256 m. These sweeping sound features are clearly not present for female calls, and are also less crucial for discriminating male calls when the entire frequency range of the signal is available. The detection of these information bearing joint spectro-temporal features is only feasible when the complete spectrographic representation of the sound (SPC parameters) is used in the discriminant function analysis.

Finally, we can also compare the DFs obtained at each distance to the single set of DFs that could be used to discriminate the caller identity irrespective of distance (bottom row of Fig. 9). The DFs obtained with the analysis taking all distances into account show that fine spectral cues are indeed important for discrimination, for both sexes; temporal patterns are also clearly visible for females, and to a lesser extent for males. Finally, one can note that while the "all-distances" DFs sample a relatively large frequency bandwidth, as seen in the DFs optimized for short distances, the weighting in these DFs optimizes the same narrower frequency range as in the DFs optimized for long distances (between 2 and 4 kHz). In this respect, the "all-distances" DFs illustrate the template of information within the spectrographic space that could be taken into account in order to achieve optimal discrimination at all distances.

Comparison between frequency power spectrum of calls and frequency gain of discriminant functions

In Fig. 4, we have shown the effect of propagation on the frequency power spectrum of the zebra finch distance call. Here we compare the frequency dependent SNR shown in those power spectra with the frequency gain of the discriminant functions: we calculated the normalized means of the absolute values of the DF for each frequency window (taking into account the first 4 DFs, weighted with their corresponding eigenvalues) and plotted this relative importance as a function of frequency. The results of that calculation are shown on Fig. 10. For close distances (2 m and 16 m), the "critical band-pass" in the signal that is of importance for individual discrimination is broad, ranging from approximately 700 Hz to 6 kHz. At further distances (64 m and 128 m) this critical band-pass is reduced mostly in the lower frequencies, ranging from approximately 2 to 6 kHz, and at 256 m the interval is reduced again, mostly in the higher frequencies this time, ranging from approximately 2 to 4 kHz. Thus, both the peaks of these DF frequency gain curves and their bandwidth match the shape of the FS of the calls with a peak approximately at 2.7 kHz for the females and 3 kHz for the males and a decrease in bandwidth as the propagation distance increases. However, one can also observe that the bandwidth of the DF frequency gain curves does not decrease as linearly as the FS of the calls: the gain remains high until the SNR drops below a certain level. This effect is clear for example for the information between 500 Hz and 2 kHz that is much higher in the DF frequency gain curves than what one would predict from the frequency spectra; similar effect is shown at higher frequencies (between 5 and 6 kHz). By examination of the DFs, one can see that this lower frequency shoulder includes spectral information on the pitch of the call (corresponding to the 2nd harmonic) that is particularly well resolved at these frequencies. The higher frequency shoulder appears to have both fine and coarse spectral information (pitch and formants). This analysis shows that the relationship between SNR and the gain of filters that can extract specific information in the signal is not proportional in the way that it would be for a simple detection task (and given, for example, by the Wiener filter for Gaussian signals). Instead, the frequency gain curve for a discrimination task must also takes into account the frequency dependent weight of the information content.

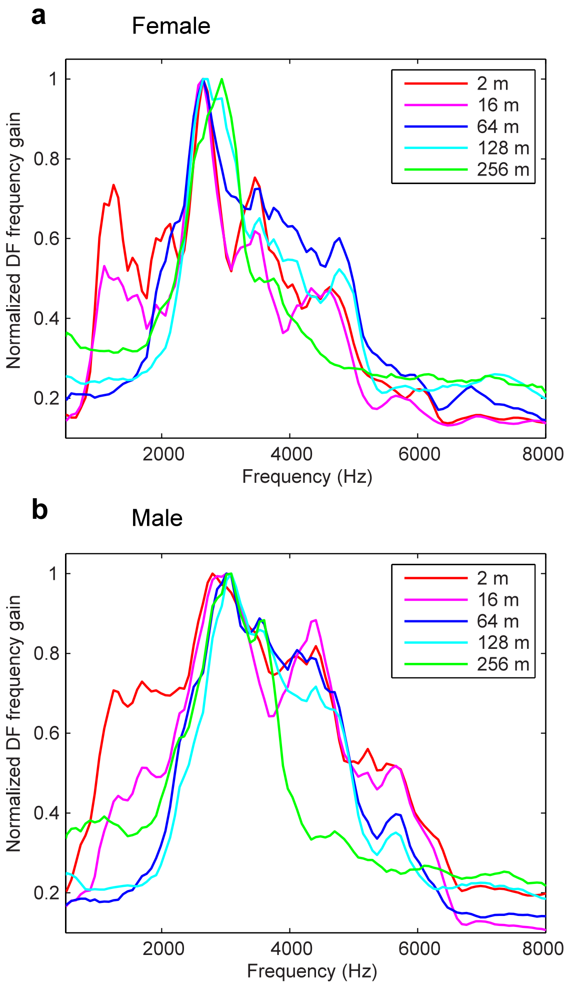


Figure 10. Relative importance of the discriminant functions (DFs) as a function of frequency. Calculations are shown for females (a) and males (b), for every propagation distance. These *importance* functions were obtained by calculating the normalized mean of the absolute values of the first 4 DFs for each frequency window. This analysis shows the frequencies that are the most useful to discriminate between individuals at each distance.

Discussion

Our analysis shows that the individual signature encoded in the distance calls of zebra finches is remarkably resistant to propagation-induced sound degradation: after more than a hundred meters of propagation, distance calls still allow for a good discrimination between individuals, and even after propagating over 256 m it remains well above chance when using acoustical features optimized for that distance. In addition, we have confirmed that information about the individual identity is carried redundantly, using multiple acoustical features [9], [21]. Moreover, the features providing the highest discrimination at long distances are not the same as those that provide the highest information at short distances. On one hand, the redundancy could therefore be used to optimize information transmission at different distances. On the other hand, we also found that a single set of acoustical features could be used to maximize information about identity irrespective of distance (except for the furthest, 256 m). Finally, whereas discriminability for both sexes is identical at long range using the parameters of the envelopes (extracted from temporal or spectral domains), male calls show a stronger individual signature at long distance compared to female calls when using the SPC parameters (extracted from the whole spectrogram of the calls). This sexual dimorphism is not due to higher signal power in the male call but to structural differences in the distance calls of each sex. These structural differences do not seem to affect the discriminability between individuals for either sex at close and medium range, but may provide an advantage to male signature over longer distances.

The results of the discriminant analysis confirm and complete our knowledge on the sexual dimorphism in the individual signatures of the zebra finch distance call. In males, the faster and more predominant downsweep component appears to become critical for discriminability at long distances. Interestingly, while the DFs for the females are comparable in structure to spectrograms of female calls, the DFs corresponding to the males are very different from the spectrographic representation of an average male call, the downsweep component in particular being absent for short and medium distances and only appearing at 128 m and 256 m. These fine spectro-temporal modulations appear therefore less critical for discrimination at short and medium distances. Information about frequency modulation is therefore not necessary at close range. The observation that the information embedded in the parameters from either the spectral or the temporal envelope is not independent (Fig. 8) further confirms this conclusion. An acoustic analysis using parameters describing the modulation of the fundamental frequency would indeed find that the downsweep component of male calls is highly individualized [9]; what our study highlights however is 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. Concerning female calls, Vignal *et al.* [8] found that a variety of spectral parameters describing the frequency modulation of the fundamental frequency were highly individualized. As shown on Fig. 7a, the higher discriminability in the female calls found for the SPC parameters compared to the envelope parameters indeed shows that information about frequency modulation is important for discrimination between females, even at short range, which is in accordance with previous work.

The frequency spectrum (FS) of our propagated calls show that the band-pass of the signals is progressively reduced with distance. The transfer function calculation shows that the frequency dependent attenuation is relatively flat in the region of interest and that the attenuation can be explained within 5 dB by spherical spreading and atmospheric absorption. Thus, although it is known that higher frequencies are more strongly attenuated through propagation-induced degradation [1] and that ground effects can cause attenuation at low frequencies, these effects appear to play a negligible role for these signals and open environment. The negligible ground effect may be explained by the source-receiver elevation of 1.3 m [5], [30]. Therefore, the resulting band-pass quality of the FS (Fig. 4) and the matched frequency gain of the DFs (Fig. 10) can be simply explained by the FS of the calls that is also band-pass limited, the FS of the noise that decreases with frequency and a relatively flat attenuation (frequency-wise) due to propagation [31]. In an environment where the spectrum of the background noise would have a significantly different shape or where the terrain and vegetation would result in a different transfer function profile, the informative frequency channel could be different. However, as described above, our analysis using recordings performed in an environment where the noise spectrum is typical of a low wind day in an open field with relatively little vegetation can be considered to be typical of the scenario encountered by the zebra finch in its native environment. Interestingly, a study using synthesized calls of zebra finches have shown that only a small number of harmonics in the distance call is necessary to elicit a behavioral response from the other sex [32]; for both males and females, this minimum bandwidth (3200 to 3500 Hz) fits inside the bandwidth containing individual information that we found at 256 m (2000 to 4000 Hz). Keeping in mind that this study did not deal with individual recognition, it is still interesting to note the convergence between information-bearing features highlighted by the acoustical analysis and signal features shown to be particularly biologically relevant.

The match between calls features and environmental characteristics has been the subject of extensive research about the influence of the acoustic properties of the environment on the selection of acoustic signals. Morton [33] suggested that the design of acoustic signals should maximize their transmission distance with regards to the characteristics of the habitat. For best resistance to degradation in an open environment, an ideal call should be short and have a high frequency modulation rate [34]. The distance calls of male zebra finches typically present these qualities, and our analysis shows that their individual signature remains stronger at long distance compared to the female calls. In this light, the fact that the females' calls are not more frequency modulated raises the question of partners' recognition. It is possible that partners rely more on the male's call, whose individual signature is more resistant to degradation, to locate each other. Indeed, Zann observed that males do call more frequently than females [19]. To investigate this hypothesis one could study the dynamics of movements when the mates reunite after being separated visually but not acoustically, and test if the females move more towards the males than the opposite, depending on the distance between mates or the level of degradation of the calls. Besides, it has been demonstrated that male zebra finches increase their vocal amplitude with increasing distance to addressed females [35]: this ability may further increase the active space of the signature embedded in male calls, and females could also increase the amplitude of their calls to compensate for their weaker individual signature at long distance.

Only playback experiments will ultimately allow testing if male calls ensure better partner's recognition than female calls at very long distances. However, adaptation to the habitat characteristics is not the only criteria in the selection of vocal signals [36–38]. As the distance call is a learned vocalization in males (contrary to the females), its sex-specific characteristics may be a by-product of song learning, rather than the result of a strong selective shaping for long distance propagation. Influences other than habitat adaptation, such as sexual selection or learning processes, may have mainly induced the acoustic differences found in the calls of both sexes.

At the other end of the transmission chain, we can now consider the perspective of a bird decoding the individual identity in a propagated signal. As described in the results, one can imagine that the decoder could use a distance dependent set of acoustical features to determine the call identity. This strategy would maximize discriminability of identity but perhaps with the cost of the additional cognitive processing required to extract a large set of high-level acoustical features. Alternatively, we have shown that a single set of features obtained to maximize the average discrimination over a range of behaviorally relevant distances performs exceptionally well. Birds may therefore adopt a strategy where only one set of acoustical features is used for all distances. Behavioral experiments, assessing the psycho-auditory abilities of birds in discriminating between various degraded acoustic signals, or using synthetic sounds where particular acoustical features are systematically manipulated, could be designed to answer these questions.

Acknowledgements

We warmly thank Colette Bouchut and Nicolas Boyer for technical support in the ENES laboratory, and Mike Schachter and Tyler Lee (UC Berkeley) for valuable remarks during the analysis process. We are grateful to the University of New South Wales for access to the Arid Zone Research Station at Fowlers Gap, as well as to Pr. Simon Griffith and Dr. Mylène Mariette (Macquarie University) for access to the zebra finch colony.

Author Contributions

Conceived and designed the experiments: SCM NM FET. Performed the experiments: SCM NM JEE CV. Analyzed the data: SCM FET. Wrote the paper: SCM. Corrected and improved the manuscript: JEE NM FET CV.

References

1. Forrest TG (1994) From sender to receiver: propagation and environmental effects on acoustic signals. Integr Comp Biol 34: 644–654.

2. Wiley RH, Richards DG (1982) Adaptations for acoustic communication in birds: sound transmission and signal detection. In: Kroodsma DE, Miller EH, editors. Acoustic Communication in Birds. New York: Academic Press, Vol. 1. pp. 131–181.

3. Brenowitz EA (1982) The active space of red-winged blackbird song. J Comp Physiol A: Neuroethology, Sensory, Neural, and Behavioral Physiology 147: 511–522.

4. 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.

5. Marten K, Marler P (1977) Sound transmission and its significance for animal vocalization. Behav Ecol Sociobiol 2: 271–290.

6. Charlton BD, Reby D, Ellis WAH, Brumm J, Fitch WT (2012) Estimating the Active Space of Male Koala Bellows: Propagation of Cues to Size and Identity in a Eucalyptus Forest. PLOS ONE 7: e45420.

7. Robertson BC (1996) Vocal mate recognition in a monogamous, flock-forming bird, the silvereye, *Zosterops lateralis*. Anim Behav 51: 303–311.

8. Vignal C, Mathevon N, Mottin S (2004) Audience drives male songbird response to partner’s voice. Nature 430: 448–451.

9. 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 Proc 77: 191–198.

10. Lovell SF (2004) Neighbor-stranger discrimination by song in a suboscine bird, the alder flycatcher, *Empidonax alnorum*. Behav Ecol 15: 799–804.

11. Naguib M, Todt D (1998) Recognition of neighbors' song in a species with large and complex song repertoires: the Thrush Nightingale. J Av Biol 29: 155–160.

12. Jouventin P, Aubin T (2002) Acoustic systems are adapted to breeding ecologies: individual recognition in nesting penguins. Anim Behav 64: 747–757.

13. Jouventin P, Aubin T, Lengagne T (1999) Finding a parent in a king penguin colony: the acoustic system of individual recognition. Anim Behav 57: 1175–1183.

14. Aubin T, Mathevon N, Silva ML, Vielliard JME, Sèbe F (2004) How a simple and stereotyped acoustic signal transmits individual information: the song of the white-browed warbler *Basileuterus leucoblepharus*. Anais da Academia brasileira de Ciencias 76: 335–344.

15. Mathevon N, Aubin T, Vielliard J, da Silva M-L, Sèbe F, et al. (2008) Singing in the rain forest: how a tropical bird song transfers information. PLOS ONE 3: e1580.

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

17. Butterfield PA (1970) The pair bond in the zebra finch. Social behaviour in birds and mammals: Essays on the social ethology of animals and man. London: J. H. Crook. pp. 249–278.

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

19. Zann R (1984) Structural variation in the zebra finch distance call. Zeit Tierpsychol 66: 328–345.

20. Vicario DS, Naqvi NH, Raksin JN (2001) Sex differences in discrimination of vocal communication signals in a songbird. Anim Behav 61: 805–817.

21. Forstmeier W, Burger C, Temnow K, Derégnaucourt S (2009) The genetic basis of zebra finch vocalizations. Evolution 63: 2114–2130.

22. Jacot A, Reers H, Forstmeier W (2010) Individual recognition and potential recognition errors in parent–offspring communication. Behav Ecol Sociobiol 64: 1515–1525.

23. Mulard H, Vignal C, Pelletier L, Blanc A, Mathevon N (2010) From preferential response to parental calls to sex-specific response to conspecific calls in juvenile zebra finches. Anim Behav 80: 189–195.

24. Griffith SC, Pryke SR, Mariette M (2008) Use of nest-boxes by the Zebra Finch (*Taeniopygia guttata*): implications for reproductive success and research. Emu 108: 311–

319.

25. Perez EC, Elie JE, Soulage CO, Soula HA, Mathevon N, et al. (2012) The acoustic expression of stress in a songbird: Does corticosterone drive isolation-induced modifications of zebra finch calls? Hormones Behav 61: 573–581.

26. Thomson D (2007) Jackknifing Multitaper Spectrum Estimates. IEEE Signal Proc Mag 24: 20–30.

27. Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am 114: 3394–3411.

28. Okanoya K, Dooling RJ (1987) Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. J Comp Psychol 101: 7–15.

29. Mathevon N, Koralek A, Weldele M, Glickman SE, Theunissen FEE (2010) What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. BMC Ecol 10: 9.

30. Cosens SE, Falls JB (1984) A comparison of sound propagation and song frequency in temperate marsh and grassland habitats. Behav Ecol Sociobiol 15: 161–170.

31. Wiley RH (1991) Associations of song properties with habitats for territorial oscine birds of Eastern North America. Am Nat 138: 973–993.

32. Vignal C, Mathevon N (2011) Effect of acoustic cue modifications on evoked vocal response to calls in zebra finches (*Taeniopygia guttata*). J Comp Psychol 125: 150–161.

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

34. Ey E, Fischer J (2009) The "acoustic adaptation hypothesis"—a review of the evidence from birds, anurans and mammals. Bioacoustics 19: 21–48.

35. Brumm H, Slater PJB (2006) Animals can vary signal amplitude with receiver distance: evidence from zebra finch song. Anim Behav 72: 699–705.

36. Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. Funct Ecol 21: 134–142.

37. Ryan MJ, Kime NM (2003) Selection on long-distance acoustic signals. In: Simmons AM, Popper AN, Fay RR, editors. Acoustic Communication. New York: Springer-Verlag. pp. 225–274.

38. Ryan MJ, Cocroft RB, Wilczynski W (1990) The role of environmental selection in intraspecific divergence of mate recognition signals in the cricket frog, *Acris crepitans*. Evolution 44: 1869–1872.

Learning to cope with degraded sounds: Female zebra finches can improve their expertise at discriminating between male voices at long distance

Solveig C. Mouterde[1,2], Julie E. Elie[2], Frédéric E. Theunissen[2], and Nicolas Mathevon[1]

[1] Equipe de Neuro-Ethologie Sensorielle ENES-CNPS CNRS UMR8195, Université de Saint-Etienne, 42023 Saint-Etienne, France

[2] Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720

Corresponding author: solveig.mouterde@gmail.com

Short title: "Vocal discrimination by female zebra finches"

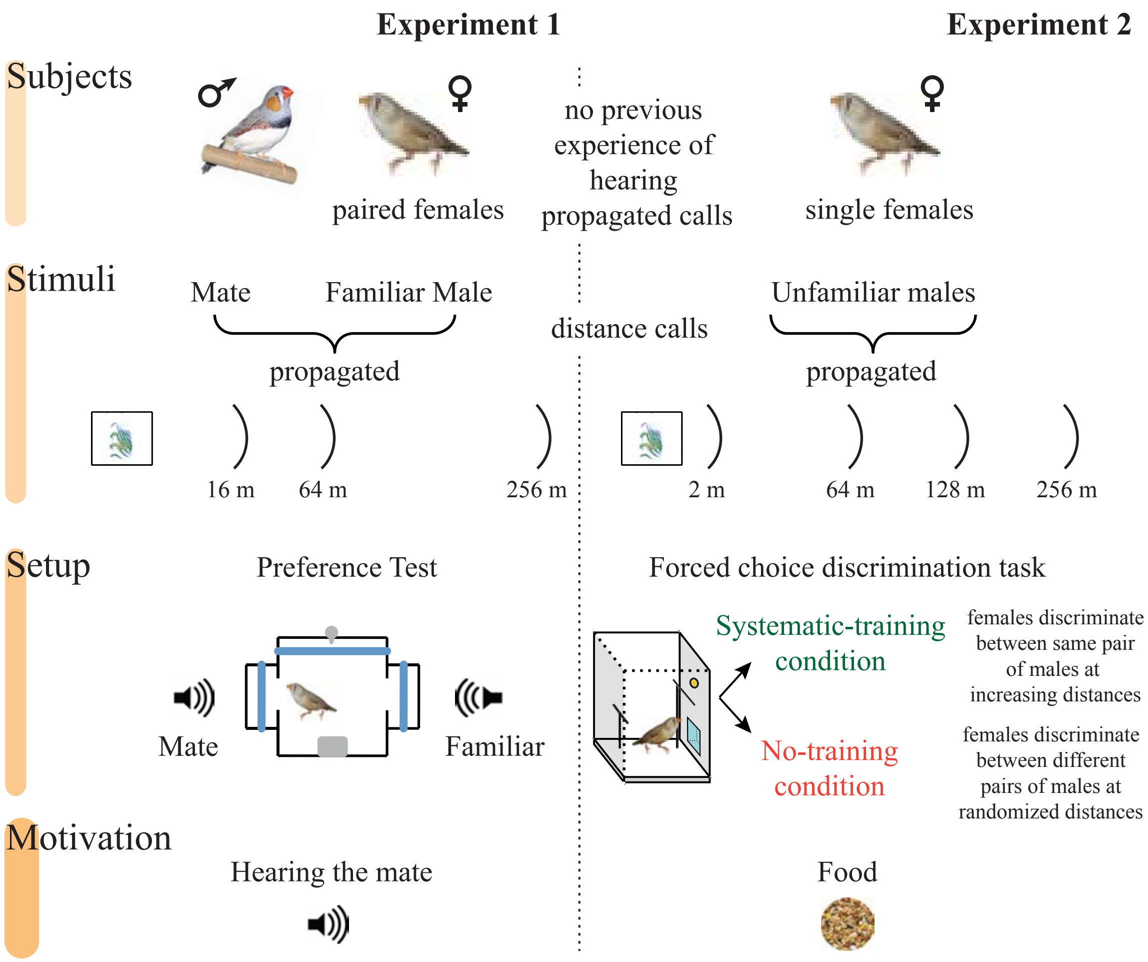
The Journal of Experimental Biology, in revision

MANUSCRIPT 2: Learning to cope with degraded sounds: Female zebra finches can improve their expertise at discriminating between male voices at long distance.

Question

How well do female zebra finches discriminate between the degraded voices of males?

Methods: Behavioral experiments



Experiment 1

Experiment 2

Subjects

paired females

no previous experience of hearing propagated calls

single females

Stimuli

Mate

Familiar Male

propagated

distance calls

Unfamiliar males

propagated

16 m

64 m

256 m

2 m

64 m

128 m

256 m

Setup

Preference Test

Mate

Familiar

Forced choice discrimination task

Systematic-training condition

females discriminate between same pair of males at increasing distances

No-training condition

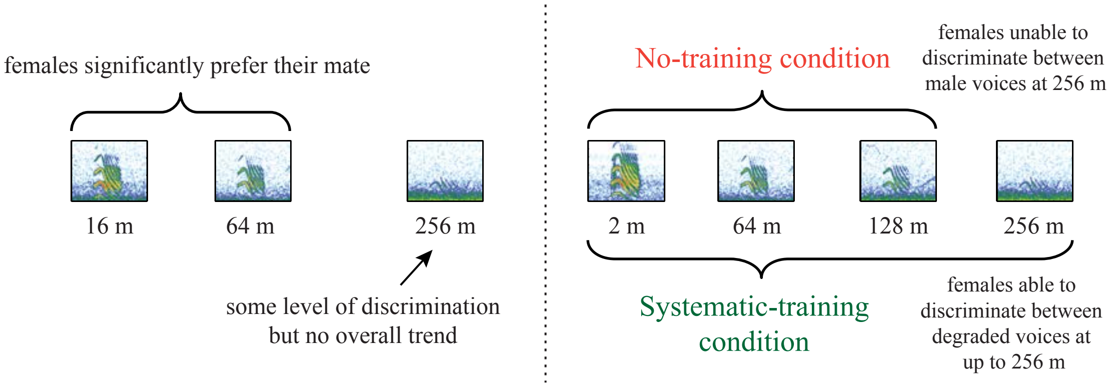
females discriminate between different pairs of males at randomized distances

Motivation

Hearing the mate

Food

Results



females significantly prefer their mate

16 m

64 m

256 m

some level of discrimination but no overall trend

No-training condition

females unable to discriminate between male voices at 256 m

2 m

64 m

128 m

256 m

Systematic-training condition

females able to discriminate between degraded voices at up to 256 m

Females are proficient at discriminating between calls degraded at long distance, and even more so when they can train themselves with increasingly degraded signals over time.

Learning to cope with degraded sounds: Female zebra finches can improve their expertise at discriminating between male voices at long distance

Solveig C. Mouterde, Julie E. Elie, Frédéric E. Theunissen and Nicolas Mathevon

ABSTRACT

Reliable transmission of acoustic information about individual identity is of critical importance for pair bond maintenance in numerous monogamous songbirds. However, information transfer can be impaired by environmental constraints such as external noise or propagation-induced degradations. Birds have been shown to use several adaptive strategies to deal with difficult signal transmission contexts. Specifically, a number of studies have suggested that vocal plasticity at the emitter's level allows birds to counteract the deleterious effects of sound degradation. Although the communication process involves both the emitter and the receiver, perceptual plasticity at the receiver's level has received little attention. Here, we explored the reliability of individual recognition by female zebra finches (*Taeniopygia guttata*), testing whether perceptual training can improve discrimination of degraded individual vocal signatures. We found that female zebra finches are proficient in discriminating between calls of individual males at long distance, and even more so when they can train themselves with increasingly degraded signals over time. In this latter context, females succeed in discriminating between males as far as 250 m. This result emphasizes that adaptation to adverse communication conditions may not only involve the emitters' vocal plasticity, but also the receptors' decoding process through ongoing learning.

KEY WORDS: Acoustic communication, Vocal recognition, Perceptual plasticity, Propagation-induced degradation, Noise, Learning

SYMBOLS AND ABBREVIATIONS:

LOR: $\log_2$ of the odds ratio

OI_{NoRe}: odds for interrupting the Non Rewarded (NoRe) stimuli

OI_{Re}: odds for interrupting the Rewarded (Re) stimuli

OR: odds ratio

GLMM: generalized linear mixed-effects model

INTRODUCTION

As monogamy represents the dominant avian mating system (Emlen and Oring, 1977) and given the importance of sound communication in birds, interactions between paired-mates based on reliable information transmission through the acoustic channel are critical to the ultimate fitness of most bird species (Falls, 1982; Kondo and Watanabe, 2009). Yet, vocal communication may be challenging due to the intrinsic nature of signal propagation and environmental noise (produced by other animals, wind, or human activity). As sounds propagate through the environment, their quality is degraded, affecting the signal's amplitude as well as its temporal and spectral structure (Forrest, 1994; Wiley and Richards, 1982). As a result, the information-bearing features in communication signals can be severely compromised, reducing the signals' active space, that is, the distance from the source over which the signal remains biologically relevant for potential receivers (Brenowitz, 1982; Marler and Slabbekoorn, 2004). How individuals cope with the environment-induced degradation of sound signals could therefore play an important role in pair-bonding birds, specifically if mates have to recognize each other by voice at long range.

Birds have been shown to be proficient in communicating in difficult listening situations (Aubin and Jouventin, 2002; Brenowitz, 1982; Klump, 1996; Park and Dooling, 1986). Individuals may alter their vocalizations, e.g. by modifying the amplitude and the pitch of songs and calls to increase the signal-to-noise ratio (Brumm, 2004; Mockford and Marshall, 2009; Nemeth et al., 2013; Slabbekoorn and Peet, 2003). Behavioural strategies such as choosing optimal emission and listening posts may also help counteract the deleterious effect of environmental constraints (Dabelsteen and Mathevon, 2002; Mathevon et al., 1996). In addition, songbirds show remarkable perceptual abilities to discriminate between noisy signals (Brémond, 1978; Hulse, 2002). However, only a small number of studies have examined individual discrimination in degraded calls (Aubin and Jouventin, 1998; Jouventin et al., 1999; Mathevon et al., 2008; Vignal et al., 2008). More specifically, only one study has investigated the question of long-range individual recognition in a songbird living in an acoustically constraining environment: in the white-browed warbler *Basileuterus leucoblepharus*, a species from the Brazilian Atlantic forest, the individual signature encoded in the male song degrades rapidly during propagation, restricting individual recognition to neighbouring territorial males (Mathevon et al., 2008). Could discrimination be improved by learning? Phillmore et al. (Phillmore et al., 2002) showed that male black-capped chickadees *Poecile atriacapilla* that learned to discriminate songs from conspecifics recorded at 5 m could transfer this knowledge to discriminate the same vocalizations mildly degraded by 25 m of propagation, suggesting that training at a given distance could improve discrimination performance at longer distances. However, the role of experience in the discrimination of individual voices degraded over a range of distances up to the limit of the active space remains unknown in songbirds. In the present paper, we investigate the ability of female zebra finches to individually identify males based on their degraded propagated vocalizations. Furthermore, we examine whether training with increasingly degraded signals can help female subjects improve their discrimination ability.

The zebra finch *Taeniopygia guttata* is a small gregarious songbird from subarid Australia that pair bonds for life and lives in large flocks in open country with a scattering of trees and bushes (Butterfield, 1970; Zann, 1996). Because these birds are opportunistic breeders living in a very unpredictable environment (Zann, 1996), maintaining strong pair bonds between breeding events while living in large fission-fusion groups is of utmost importance. With the flock being constantly on the move, topographic landmarks may be scarce and partners might not have a fixed nest site to meet each other: using a solid vocal recognition system could avoid partners the cost of losing each other. Of all the zebra finch vocalizations, the distance call is used specifically between pair-bonded partners while foraging and when they loose visual contact (Zann, 1996). Previous studies have shown that the distance call bears an individual signature, and that the birds are capable of call-based recognition (Vignal et al., 2004; Vignal et al., 2008; Zann, 1984). The active space of zebra finch distance call has been estimated based on naturalistic observations to extend up to 100 meters (Zann, 1996); similar conclusions were reached using discrimination thresholds for masked signals in this species (Lohr et al., 2003).

In the present study, we predicted that zebra finches have a robust vocal recognition system that performs well even for distance calls propagated over long distances (i.e., longer than 100 meters). Specifically, we hypothesized that females could improve their ability to discriminate between male voices through experience. To test this hypothesis, we first explored the reliability of mate recognition by females at a range of propagation distances by assessing their preference for their mate's calls, using an operant choice apparatus (Experiment 1). Then, to assess the role of experience and to further distinguish the discrimination process from the recognition process, we conducted forced-choice conditioning experiments (Experiment 2) and compared the results of two different protocols. In both protocols of Experiment 2, the females were asked to discriminate between the calls of two unfamiliar males. These calls had been propagated over the same distance: short, medium or long range. In the first protocol ("systematic-training" condition), the females learned to discriminate the 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 ("no-training" condition), the females 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. To the best of our knowledge, this is the first study to examine the importance of learning in improving discrimination of individual vocal signatures following strong environment-induced degradation.

RESULTS

Signal degradation and the difficulty of individual discrimination at long ranges

To illustrate the increasing similarity between calls with propagation distance, we calculated the spectral correlation between the distance calls of males for every different pair that was used in both experiments. As one could expect, these correlation values are highly correlated with propagation distance ($r = 0.73$, $p < 0.001$), increasing to values close to 1 at 256 m (Fig. 1) where very little individual information in the degraded signal remains. This basic spectral analysis serves to quantify the increasing difficulty that subjects encounter when discriminating between male individuals at increasing distances. As illustrated by the spectrograms from the same pair of males displayed as an example on Fig. 1, the progressive decrease in signal-to-noise ratio at long distances results in signals that are dominated by noise and have therefore very similar frequency spectra yielding high correlation values.

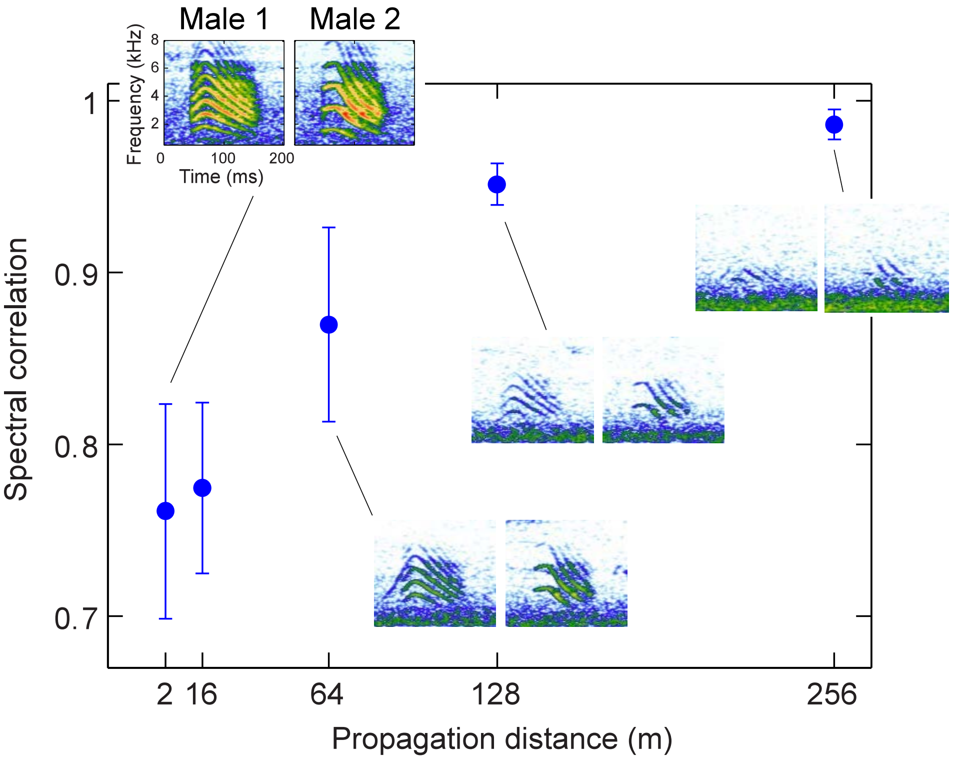


Fig. 1. Spectral correlations between the distance calls of males of every different pair to be discriminated by tested females in both experiments, as a function of propagation distance. Mean correlations are represented for each distance, and error bars correspond to the standard deviation. Correlations were calculated between the mean frequency spectra of each male. The correlation between male calls increases with distance along with the propagation-induced signal degradation and the decrease of the signal-to-noise ratio. As an example, the spectrograms of the same calls from two males used as a Re-NoRe pair are shown for each distance tested in Experiment 2.

Experiment 1: Preference test

The purpose of this experiment was to assess the unreinforced preference of female zebra finches for variously degraded distance calls from their mate or from a familiar male (propagated at 16, 64 and 256 m), using an operant choice task with call playbacks as a reward. In the experimental apparatus (see methods), the subject could trigger the playback of a degraded distance call from either its mate (Mate) or a familiar male (Non-Mate) by perching on one or the other of two perches placed in opposite sides of the cage. Each subject was tested 3 times (for 3 distances), and each test consisted of two sessions between which the assignment of the Mate and Non-Mate calls to each side was swapped. We hypothesized that females would prefer their mate's call, providing further evidence for individual recognition in zebra finches.

Using perching events in the side arms (perching right/left) as a dependent variable, we found an effect of the side assignment of Mate and Non-Mate calls (logistic regression calculated across all subjects and all distances: $\chi^2(3) = 249$, $p < 10^{-4}$), and an effect of distance ($\chi^2(9) = 3780$, $p < 10^{-4}$). Females thus expressed a differential response to the Mate side, and their perching probability was influenced by propagation distance. Post-hoc tests showed a significant preference for the Mate side at 16 m and 64 m, but not at 256 m (16 m: $\chi^2(2) = 6.01$, $p = 0.049$; 64 m: $\chi^2(2) = 8.74$, $p = 0.013$; 256 m: $\chi^2(2) = 0.28$, $p = 0.87$). We also noticed that session order had a significant effect on females' choice (16 m: $\chi^2(4) = 587$; 64 m: $\chi^2(4) = 38$; 256 m: $\chi^2(4) = 74$; all $p < 10^{-4}$). The effect of session order can be explained by the fact that birds seemed to persevere in their preference behaviour from the first session to the second session. Indeed, while the subjects could assess the side of their Mate prior to the first session (see methods), they had to get used to the side reversal during the second session, which could result in certain "inertia" in changing the side preference, even when recognition occurred.

We estimated the Mate effect for each bird and each distance independently using an odds ratio (OR) describing the preference for the Mate side (probability of perching on one side when that side broadcasts Mate calls divided by the probability of perching on the same side when it broadcasts Non-Mate calls, see methods): The log of the OR (LOR) is often used and is what we chose to display on Fig. 2. An OR greater than 1 (or LOR > 0) indicates a preference of the female for Mate calls while an OR less than 1 (or LOR < 0) indicates a preference for Non-Mate calls. For no preference, the OR will not be significantly different from 1 (and the LOR not significantly different from 0). As shown on Fig. 2, at 16 m the LOR for all but one of the tested females was greater than 0 (significant preference for Mate: 5/10 subjects; for Non-Mate: 1/10 subject). At 64 m, 7 females showed a LOR greater than 0 (significant preference for Mate: 5/10 subjects), while 3 showed a LOR less than 0 (all non significant). At 256 m, half of the females showed a LOR greater than 0 (significant preference for Mate: 2/10 subjects; significant preference for Non-Mate: 2/10 subjects). Moreover, one can observe that the LOR confidence intervals at the longer distance were larger due to a reduced amount of total perching events, suggesting that male voices were less salient for the females. Preference became more arbitrary; while a number of subjects showed a significant preference for their Mate's calls at short and medium distances, the females' choice became more randomly distributed at 256 m with a few subjects significantly preferring the Non-Mate calls.

Finally, in order to further investigate the effect of propagation distance on Mate preference, we estimated the Mate effect as explained earlier but taking into account only the 5 subjects that showed a significant preference for their Mate at 16 m. In this case we also found an effect of the Mate side ($\chi^2(3) = 262$, $p < 10^{-4}$) and an effect of distance ($\chi^2(9) = 872$, $p < 10^{-4}$). In Fig. 2, a bold black circle signals each of these 5 subjects, and the insert shows the mean LOR for each distance. The effect of distance on the subjects' preference is visible, the LOR decreasing with increasing distance. The fact that the LOR is significant for Non-Mate at 256 m is due to the effect of the only subject that showed a significant preference for the Non-Mate, the high number of perching events for this subject leading the overall estimate for all birds to be significant.

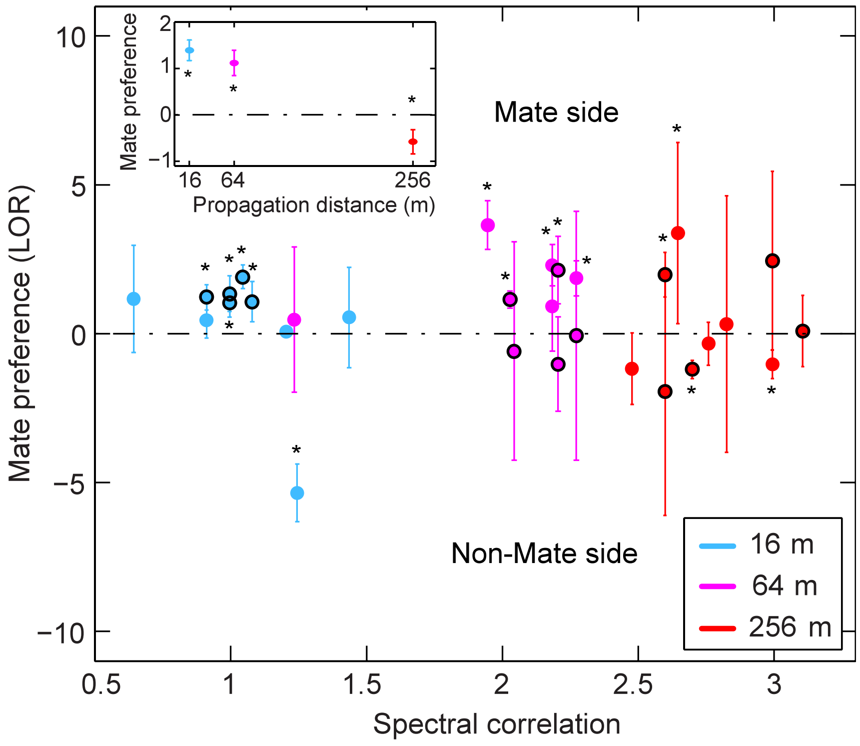


Figure. 2. Results of the females' preference tests between their Mate's and a familiar male's calls having experienced different levels of propagation-induced degradation (Experiment 1; $n = 10$ females). The figure shows the Mate preference estimates quantified by the $\log_2$ of the odds ratio (LOR) of correct choice. The LOR were estimated using a generalized linear model for each distance and each female (see main text). Positive values express a preference for the Mate's voice. For visual clarity, the spectral correlations on the x-axis are Fisher transformed to obtain an unbounded correlation measure. The 5 subjects that showed a significant preference for their Mate at 16 m are signalled for all distances by bold black circles. The insert (top left) shows the LOR of Mate preference using these 5 subjects. * $p < 0.05$. Error bars correspond to the 95% confidence intervals.

Experiment 2: Discrimination task

The purpose of this second experiment was, first, to further investigate the ability of female zebra finches to discriminate between two males while hearing variously degraded distance calls and, second, to test whether females could improve their discrimination through learning. Using a pecking key apparatus and a forced choice operant procedure, we first assessed this ability in a “systematic-training” paradigm: the tested females were asked to discriminate between the same pair of male calls from one day to the next, with increasing propagation distances (2, 64, 128 and 256 m). We compared this to a “no-training” control condition: here, both the identity of males and the propagation distances were randomized over the 4 testing days. The birds were performing a simple task: they were triggering the playback of calls at will by pecking on a key. At any time they could choose to attend the full duration of the stimulus or peck again to interrupt the current stimulus and trigger the next one. The access to the feeder was only permitted when the bird chose to fully attend to the rewarded stimuli. The subjects were thus tested on their ability to interrupt the Non Rewarded (NoRe) stimuli and refrain from interrupting the Rewarded (Re) stimuli, this behaviour ensuring them a greater access to food (see methods).

For both protocols, we retrieved for each pecking event the stimulus type (Re/NoRe calls) and the subject's response (Interruption/Non-interruption). We first assessed the overall effect of stimulus type, distance and spectral correlation between the Re and NoRe sound stimuli by modelling the interruption behaviour using logistic regression (see methods). We performed these analyses for the “systematic-training” and “no-training” conditions separately. For both protocols we found: 1) that the stimulus type (Re/NoRe) was significant, indicating that the birds were learning the task (for “systematic-training”: $\chi^2(5) = 120, p < 10^{-4}$; for “no-training”: $\chi^2(5) = 145, p < 10^{-4}$), 2) that distance was significant, indicating that the performance in the task varied as a function of distance (for “systematic-training”: $\chi^2(6) = 27$, $p = 0.0002$; for “no-training”: $\chi^2(6) = 49, p < 10^{-4}$) and 3) that correlation between sound stimuli was significant, indicating that the task performance was affected by the degree of similarity between the two sounds (for “systematic-training”: $\chi^2(2) = 55, p < 10^{-4}$; for “no-training”: $\chi^2(2) = 19, p < 10^{-4}$).

To visualize the effect of distance on this discrimination task and to better analyse the differences between the two paradigms, we calculated, across subjects, the probability of interrupting each stimulus type (Re and NoRe) at each distance and for each paradigm separately (see methods). Statistical significance was assessed using logistic regression (see methods). In the “systematic-training” condition, the probability for interrupting the NoRe stimuli (P_{NoRe}) were significantly higher than the probability for interrupting the Re stimuli (P_{Re}) for all distances including 256 meters (2 m: $\chi^2(1) = 14.6$, $p = 0.00013$; 64 m: $\chi^2(1) = 25.7$, $p < 10^{-4}$; 128 m: $\chi^2(1) = 39.8$, $p < 10^{-4}$; 256 m: $\chi^2(1) = 33.5$, $p < 10^{-4}$; Fig. 3A). These results indicate that the subjects were able to discriminate between the Re and NoRe stimuli at up to 256 m. The odds ratio, defined here as the odds of interrupting the NoRe divided by the odds of interrupting the Re (OI_{NoRe} / OI_{Re}) can then be used to estimate the effect size of the differences in the probability of interruption. These odd ratios were relatively constant across distances (2 m: OR = 1.83; 64 m: OR = 2.23; 128 m OR = 2.1; 256 m OR = 2.2), showing that, apart from the slight improvement from 2m to 16 m as birds learned the task, their performance remained constant as the distance increased although the task became more difficult. This preservation of performance appears to result from the experience gained in previous training days with the easier task.

Conversely, for the ”no-training” condition, the probability of interrupting the NoRe (P_{NoRe}) were significantly higher than the probability for interrupting the Re stimuli (P_{Re}) at 2, 64 and 128 m but not at 256 m (2 m: $\chi^2(1) = 17.8$, $p < 10^{-4}$; 64 m: $\chi^2(1) = 99.8$, $p < 10^{-4}$; 128 m: $\chi^2(1) = 10.4$, $p = 0.0012$; 256 m: $\chi^2(1) = 0.28$, $p = 0.6$; Fig. 3B). Females were thus not able to discriminate between the calls of two males at this longer distance. In this case the effect of distance is also reflected in the odds ratio (OI_{NoRe} / OI_{Re}) that tends towards 1 for the longer distances (2 m: odds ratio or OR = 2.0; 64 m: OR = 6.4; 128 m OR = 1.7; 256 m OR = 0.93). Without training the task appears thus more difficult at longer distances.

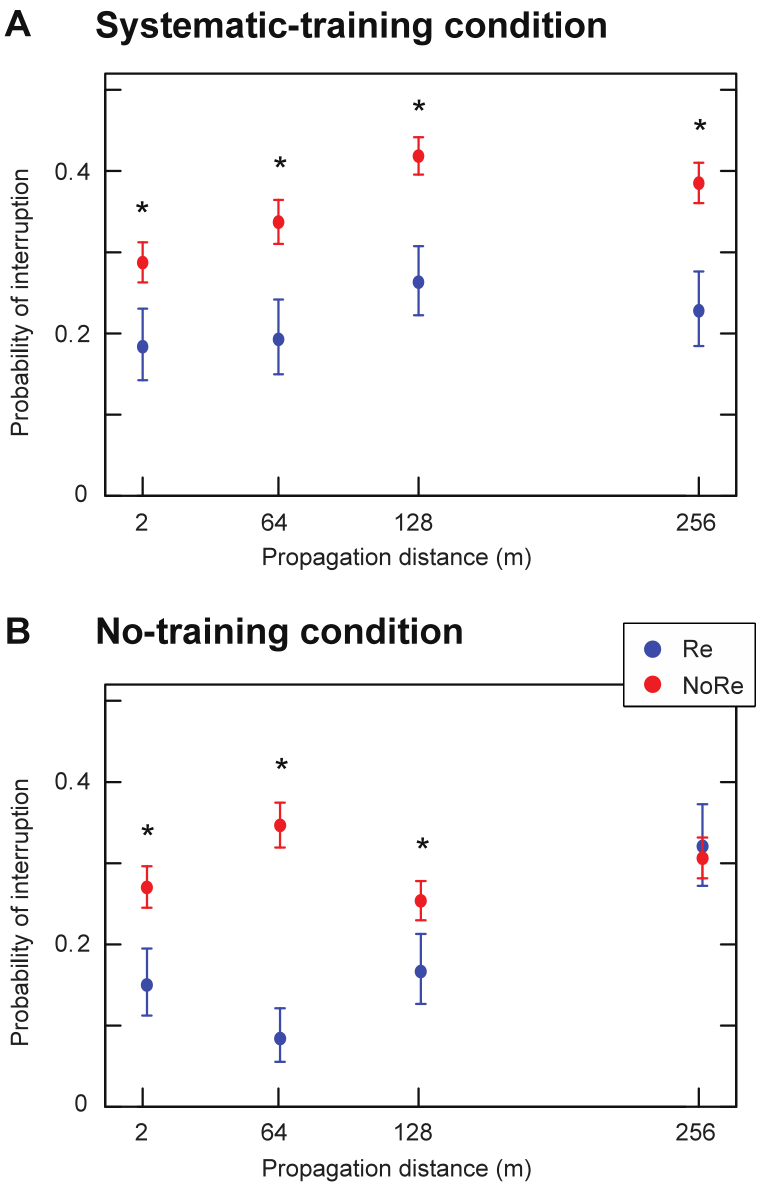


Figure. 3. Results of the discrimination tasks between calls of two different males performed by females (Experiment 2). (A) Results for the “systematic-training" condition (the same pair of males was used over all distances; females were challenged with increasing distances; n = 5). (B) Results for the “no-training" condition (pairs of males changed for each distance; distances presented at random; n = 7). The figure shows the average probability of interrupting the Re and NoRe stimuli for each tested propagation distance (see methods). Significantly higher odds values for interrupting the NoRe stimuli compared to the Re stimuli indicate that female subjects have been able to discriminate between the two sets of stimuli and responded accordingly to get access to food. Error bars represent 95% confidence intervals on the probabilities (binomial test). * shows the significance obtained from the logistic regression ($p < 0.002$).

DISCUSSION

Using two complementary approaches, we showed that female zebra finches are proficient in discriminating between the calls of two individual males at long distances, and even more so when they have the possibility to learn over time. This ability may be highly adaptive for this life-long monogamous species, as losing the partner would represent a high cost due to their unpredictable living environment. Our result underlines the importance of considering the receiver's level when studying acoustic communication in adverse conditions.

In the first approach, we aimed to determine if female zebra finches were capable of discriminating between the calls of their mate and of a familiar male at different distances. The preference test showed that females significantly prefer the call of their mate at 16 m and 64 m, but not at 256 m, where the estimates were evenly distributed around zero. The fact that a few subjects significantly preferred the Non-Mate calls at this latter distance could be explained by the fact that the subjects were not able to discriminate the call of their mate but still managed to detect differences in the two playbacks and showed a significant preference for one (Lohr et al., 2003). With our second approach we further investigated long distance discrimination and assessed the zebra finches' discrimination abilities per se, disentangling the subjects' recognition process from the preference for their mate's calls and eliminating the potential impact of any previous social interactions with the males used as stimuli in the experiment. The conditioning experiment showed that female zebra finches are indeed able to discriminate between two male individuals at up to 128 m in a "no-training" context, but that when they have the possibility to learn from their previous experience, robust discrimination occurs even as far as 256 m.

In the field (Zann, 1996), the active space of zebra finches' distance calls has been estimated to be up to 100 m. A similar measure has been obtained based on discrimination tasks at different signal to noise ratio performed in the laboratory (Lohr et al., 2003). Here we show that, with training and also in a laboratory setting, female zebra finches have the physiological ability to recognize acoustic signatures at least up to 256 m. One has to be careful of course when comparing results in the field and in the laboratory. On the one hand, laboratory experiments have the advantage of being reproducible and allowing a precise quantification of behavioural discrimination. On the other hand, by design laboratory conditions tend to focus on a single or a few stimulus parameters (e.g. the signal to noise) and miss other factors that will be found in the more complex natural conditions. For example, in nature, birds might encounter more adverse propagation conditions (e.g. strong winds) or interference from other sound sources including other conspecifics. Nonetheless, our results suggest that wild zebra finches could benefit from repeated learning experiences in the recognition of their partner's degraded individual signature and achieve discrimination at distances greater than 100. This enhanced learned discrimination might also come into play in more adverse conditions, preserving discrimination at shorter distances. One should also note that laboratory experiments might underestimate the natural discriminability. For example in our experiments, the signal, the echoes from reverberation and the noise all came from the same spatial location: the loudspeaker. In nature, these different sounds can come from spatially separated sound sources and spatial information could thus be used to further enhance discrimination (Bee, 2008; Dent et al., 2009; Maddox et al., 2012). A full recognition task in the natural context might also combine the tasks of discriminating one particular individual and determining its spatial location, i.e. its orientation and distance. Previous laboratory experiments showed that zebra finches can also discriminate between near and far signals (Phillmore et al., 1998; Radziwon et al., 2011) and have a rough sense of sound source azimuth (Park and Dooling, 1991). Further experiments both in the field and in the laboratory are needed to assess the performance for both localization and identification and in more complex listening conditions.

In light of what could be considered a remarkable performance for individual discrimination in zebra finches, it is interesting to compare these measures to those obtained in other species. For the territorial white-browed warbler living in the dense environment of the tropical forest, the male’s song loses its individual signature after less than 100 m of propagation; thus, the transmission of individual information is likely to be limited to nearby individuals, i.e. territorial male neighbours and the female partner (Mathevon et al., 2008). In the context of pair bond, females in this species would have access to the individual signature of their mate within the limits of their territory. Another interesting example is given by female great tits incubating inside nest boxes that still perceive subtle individual differences between their mate’s song and a neighbour’s song emitted from outside the box, despite the similarity between the songs and the sound-degradation induced by the nest box (Blumenrath et al., 2007). In this case, females have to deal with short-range signal degradation but with a difficult discrimination task due to the signal similarity between individuals. Conversely, the ecological requirements of the zebra finch, especially its nomadic lifestyle outside of the breeding events in an open environment, may call for perceptual abilities in individual recognition that are adapted to vocalizations being propagated at longer distances. Although male zebra finch calls have been shown to be clearly distinct acoustically (Zann, 1984), the acoustical discriminability of the individual signature in communication calls across different songbird species occupying different ecological niches remains an open question.

In our design, we chose to test females as they were shown to respond preferentially to their mate's voice (Miller, 1979; Vignal et al., 2008), whereas males' responses can change depending on the composition of the audience (Vignal et al., 2004). Testing males in the same tasks would be interesting; although it is not certain that we could reliably test the males' abilities to recognize degraded calls using their preference for their partner's calls in an isolation context, using a conditioning experiment for testing male discrimination would certainly be insightful. Indeed, in zebra finches the individual signature was found to be stronger in male distance calls than in the females' (Zann, 1984); discriminating between degraded calls of females could therefore be more difficult than for male calls.

In the complex task of recognizing individual voices in propagation-induced degraded calls, another interesting question is to assess the extent to which learning takes part in the recognition process. Using field-reared and isolate-reared songbirds (*Poecile atriacapillus*), Phillmore *et al.* (Phillmore et al., 2003) showed that the discrimination of distance cues (i.e., the emitter's perception of its distance to the sender) was probably an innate skill. Conversely, the recognition of individual vocalizations appeared to require auditory contact with adult conspecifics during the subjects' development. In the combined task of extracting information about individual identity in degraded calls, our indoor colony-reared subjects showed impressive abilities in the no-training context, and even greater capacities when given the opportunity to learn from one day to the next. These results suggest that the zebra finches' vocal recognition system is highly efficient for degraded calls, and that it can be further improved through perceptual plasticity.

Further studies investigating how learning influences this laboratory task, using synthesized stimuli controlling for the acoustic similarity between calls and the level of call degradation, would be useful for a better understanding of the ongoing plasticity of the auditory discrimination process in adult songbirds. Besides, studying plasticity at the level of the individual, as a means to adapt quickly to varying environmental conditions, would be of primary importance in the current context of ever-growing anthropogenic noise. While the vocal plasticity of the sender has been the subject of a number of recent studies (Francis et al., 2010; Nemeth et al., 2013; Warren et al., 2006), the receptor's perceptual plasticity has received much less attention (Pohl et al., 2012; Slabbekoorn, 2013). The present study emphasizes that the adaptation to adverse communication conditions may also involve ongoing learning at the receiver’s level.

MATERIALS AND METHODS

Experiment 1: Preference test

Subjects

The subjects (n = 10 adult male-female pairs) were raised in the ENES laboratory (14L/10D photoperiod with adapted wavelengths; food and water *ad libitum*; temperature between 23 and 25°C). Prior to the experiments, the pairs were observed over a 2 months period of time to assess if they were effectively mated. Every pair had thus been observed allopreening, building a nest and incubating eggs. The pairs were housed in separate cages (38 cm W x 24 cm D x 40 cm H) in the same room, having visual and vocal contact with each other.

Recording the distance calls and preparation of the stimuli

To promote calling behaviour, the male and female of each pair were kept in two separate cages and placed out of sight in two connected soundproof rooms. The males were recorded using a microphone (Sennheiser MD-42, Wedemark, Germany) placed 0.2 m above the cage and connected to a Marantz Professional Solid state recorder (PMD 670, Eindhoven, The Netherlands; sampling frequency: 44100 Hz). Conditions of temperature, food and water availability were the same as in the aviary.

We isolated 10 distance calls from each male and normalized them by matching the maximum values of the sound pressure waveforms. These calls were used to create our propagated calls database. The propagated recordings were performed on an open flat field (Bellegarde-en-Forez, France; on the 1st of March 2011 around noon, with cloudy weather and wind < 5 km/h, temperature = 10°C). We have shown that the effects of propagation on this French site are similar to those in the Australian desert with little vegetation (Mouterde et al., submitted). All 10 calls of each male bird were dispatched along an audio sequence of 4 min long, in order to avoid any context effect (e.g. changes in the background noise). The calls sequence was broadcasted from a Marantz Professional Solid state recorder/player (PMD671) connected to a MegaVox speaker (PB-35W, Anchor Audio Inc., Torrance, CA, USA) placed on a stool, 1.30 m high so as to avoid excessive ground reflection interference. The volume of the Marantz player was set to obtain mean sound level of 70 dB SPL at 1 m (Velleman Sound Level Meter DVM 1326, Gavere, Belgium) to match typical levels of the natural distance call in the zebra finch (Vignal et al., 2008). The call sequences were then recorded with a microphone (Schoeps MK4 cardioid, on a CMC6-U base, Karlsruhe, Germany) equipped with a Schoeps Basket-type Windscreen (W20) and set 1.30 m high. The microphone was connected to a second Marantz recorder/player (PMD671; sampling frequency: 44100 Hz). We recorded the calls sequence 1 m (for calibration), 16 m, 64 m and 256 m away from the source, three to four times for each distance, enabling us to select from these three or four versions of each propagated call a signal that had not been impaired by unexpected transient sounds (e.g. birds or other animals calling in the vicinity, human-related activity). From these recordings, we thus isolated 10 different calls per male per propagation distance (10 calls * 10 males * 3 distances; total = 300 calls). The background noise immediately preceding and following each call was replaced by silence; the call was then ramped (relative amplitude gradually increased from silence to full volume over 10 ms using Goldwave©) to avoid any switching noise at onset. To further remove irrelevant background noise, we also applied a high-pass filter above 500 Hz on the signals, following the lower frequency threshold of the zebra finch's audiogram (Okanoya and Dooling, 1987).

Estimating calls acoustic similarity

To evaluate the difficulty of the discrimination task, we estimated the intrinsic similarity between the distance calls of males of every pair to be discriminated in the experiment by calculating the correlation between the mean frequency spectra of all sets of calls for each pair of males, at each distance (meanspec function, Seewave R package (Sueur et al., 2008)).

Experiment setup and protocol

Preference tests were carried out in a sound attenuation chamber (internal dimensions: 1.8 m W x 1.4 m D x 2.2 m H; Silence-Box, Tip-Top Wood, Saint-Etienne, France). Each female was housed alone in the chamber, in an experimental cage with a central body (30 cm W x 34 cm D x 34 cm H) where food and water were distributed *ad libitum* and which contained a single perch (see Fig. S1 in supplementary material). On each side of the cage, an opening (10 cm W x 10 cm H) led to a side arm (20 cm W x 10 cm D x 26 cm H) containing a perch and was equipped with infrared sensors that monitored when the bird entered the arm. Custom-made software was used to monitor the subject's activity on the perches and trigger playbacks as follows: a hop on a side perch broke the infrared beam and triggered the playback of a call from a loudspeaker (Bravo Allroom, Audio Pro, Sweden) placed 20 cm away from the same side arm. Depending on the side arm, this call was randomly selected either from the 10 calls available for the tested female's mate or from the 10 calls of a familiar male. Sound stimuli were broadcasted by either of the two loudspeakers connected to an amplifier (Yamaha Natural Sound Stereo Amplifier, AX-396, Hamamatsu, Shizuoka, Japan) and a laptop. We calibrated the intensity of the sound stimuli by setting it at 70 dB SPL for the sounds recorded at 1 m (typical level of a natural distance call) and used that gain setting for all playbacks. Thus, stimuli used for the experiment (from 16 to 256 m) were emitted at the lower intensity level that matched the amplitude loss due to natural propagation.

The 3 propagation distances we tested (16 m, 64 m and 256 m) were representative of short, medium and long range. Each propagation distance was tested over a 3 day long trial and the female could choose between triggering either its mate's calls (Mate) or calls from a familiar male (Non-Mate) recorded at the same distance. The identity of the familiar male was the same for the 3 trials of a given female, but different between females. The order in which the propagation distances were tested for each bird was randomized across all subjects. The delay between the end of a trial and the beginning of the next trial for each bird was of 21 days minimum. Each trial consisted of two experimental sessions (1st session: Mate's calls emitted from one side and Non-Mate's calls from the opposite side; 2nd session: reversed positions) and started with a habituation period, enabling the subject to get used to the setup and learn which side arm was associated with which individual's calls (mate or familiar male) for the first session. Each experimental session lasted 17 hours (1.5 days, spread on two consecutive days, each session being interrupted by the 10 hours-long night time during which playbacks were turned off), which insured that the subject's activity was recorded during the same amount of time for each Mate/Non-Mate side assignment. The order of the side assignments was balanced across trials for each subject. After the end of the trial, the subject was returned to its mate in the colony room. The experimental protocol was approved by the Jean Monnet University's Animal Care Committee (authorization n°42-218-0901-38 SV 09 to the ENES lab).

Data analysis

The perching events in the side arms were analysed as a binary response variable (perch right/left) using a series of logistic regression analyses. We first used a generalized linear mixed-effects model (GLMM) to test the main effect of the Mate side (left or right) and the effect of distance on the females' perch choice (perch on the right or left arm). A random effect was used to control for the birds' potential bias for a particular cage side. We then estimated a GLMM for each propagation distance separately (16, 64 and 256 m) in order to examine the effect size and significance of Mate side for each distance; for these models we also analysed the order effect of each session. We also performed a statistical test for each subject, which allowed us to examine the results un-weighted by the average number of perching of each bird. Models were fitted using the lmer or the glm functions of R (v. 2.13.1, R Foundation for Statistical Computing).

The effect size of the presence of the Mate, as assessed by the model, can be expressed by the odds ratio (OR), i.e. the ratio of the odds of perching on one side when this side broadcasts Mate calls divided by the odds of perching on the same side when it broadcasts Non-Mate calls (the value of this odds ratio is right/left symmetrical and can be obtained either from the perches on the right or on the left). The higher the OR, the higher the female preference for her mate. In Fig. 2, we plotted the log of the OR (LOR) obtained for each bird and each distance using the output of the GLMM model. Error bars were obtained from the standard errors estimates of the regression coefficients obtained in the model fits. Finally, in order to clearly visualize the effect of distance on the subjects' preference, we also estimated the OR of Mate preference at each distance using only the subset of females that showed a significant preference for their Mate at 16 m ($n = 5$).

Experiment 2: Discrimination task

Subjects

Seven unpaired adult female zebra finches were used in this experiment. They were housed in the same single-sex cage at UC Berkeley’s animal facilities (12L/12D photoperiod, temperature between 22 and 24°C, food and water *ad libitum)*. All experimental procedures were approved by the Animal Care and Use Committee of UC Berkeley. Prior to the experiments, all subjects had previously been trained on the pecking test device and were familiar with the forced-choice procedure. The shaping sessions lasted for less than a week and two songs from different male zebra finches were used as Rewarded (Re) and Non Rewarded (NoRe) stimuli. Subjects were considered as shaped for the task when they were pecking at least 50 times per day and were interrupting the NoRe stimuli al least 20% more often than the Re Stimuli. For every subject, the experiment started on Day 0 with a shaping test, using these same two songs as stimuli. This ensured that each subject started the experiments with the same just-prior experience with the apparatus, and having heard stimuli that were different from those used for the actual experiment.

Recording of the distance calls and preparation of the stimuli

To prepare the stimuli for these experiments we used a distance calls database recorded between 2007 and 2008 from unpaired male zebra finches raised in the ENES laboratory. The recording procedure was similar to Experiment 1, with the difference that here each male bird was recorded in the presence of two females placed 3 m away and used as an audience to minimize stress, while in experiment 1 the male birds were in acoustic contact with the female mate only. This database was composed of 16 different calls exemplars from 16 different males (16 x 16 = 256 calls).

We recorded the propagated calls of this database on the 3rd of October 2010 in the afternoon, at the same location as explained above and using the same equipment (weather cloudy; no wind; temperature = 11°C). We recorded the propagated calls 1 m (for calibration), 2 m, 64 m, 128 m and 256 m away from the source, twice for each distance, and processed the recorded calls (16 calls * 16 males * 4 distances; total 1024 calls) as explained for Experiment 1.

Each acoustic stimulus used during the forced-choice discrimination task consisted of a sequence of 6 distance calls randomly selected from the 16 available calls of the same male individual for the same distance, and randomly distributed within a 6 s window.

Experimental apparatus

The forced-choice task apparatus (see Fig. S2 in supplementary material) consisted of a modular test chamber (interior dimensions 31 cm x 24 cm x 29 cm; Med Associates Inc, St. Albans, VT, USA) placed in a soundproof booth (Acoustic Systems, MSR West, Louisville, CO, USA; interior dimensions 76 cm x 61 cm x 49 cm). The experimental panel consisted of a pecking key placed 20.5 cm high from the floor and accessible through a wooden perch. Below, a feeder containing seeds could be made accessible or not to the subject, depending on its appropriate response to the playback. Acoustic stimuli were broadcasted by a computer connected to an amplifier (Technics, Matsushita Electronics SA-EX140, Osaka, Japan) and a loudspeaker (Bose model 141, Framingham, MA, USA) placed 20 cm from the test chamber (sound level calibrated as in Experiment 1 to match the natural intensity levels at each propagation distance). The computer was also connected to the test chamber apparatus to record pecking events, play sounds and activate the feeder in real-time with a single customised program written in Matlab©.

Conditioning procedure

Every male call used for the playbacks was unknown to the female subjects prior to the experiments. Both protocols ("systematic-training" and "no-training") consisted of 4 tests conducted for 4 consecutive days. One test consisted of three 30 min trials separated by two 90 min-long rest periods. The pecking key's light was used to distinguish the trial period (pecking light on) from the rest period (pecking light off). The 30 min countdown for each trial started when the subject pecked the key for the first time. When pecking the key during a trial (see Fig. S3 in supplementary material), the female triggered the playback of calls from either of two males randomly selected from our database: the Re male (with a probability of 0.2) or the NoRe male (with a probability of 0.8). She could then go to the feeder and wait until the end of the 6 s playback to get a reward for the Re stimuli, or interrupt it by pecking again to trigger a new stimulus. Because the time windows for pecking (the three 30 minutes trials) were limited and most stimuli were NoRe, the subjects were motivated to interrupt the NoRe stimuli until they obtained a Re stimulus, at which point waiting until the end of the playback would ensure them access to seeds for 10 s. Interrupting the playback of a Re stimulus eliminated the possibility of reward following this playback.

To motivate the subjects to use the pecking key for food reward, they were fasted 20 hours prior to the beginning of the experiment and maintained in a fasted state (85-90% of their free feeding weight) for the whole experiment by only giving them 1.5 g of seeds per individual after each daily test. Every day, the subjects' weight was monitored before starting the test and was returned to its cage in the colony room after the test. As approved by the Animal Care and Use Committee of UC Berkeley, our criterion for interrupting the fast was a loss of weight superior to 15% of the initial weight recorded before the fasting started. No bird was taken out of the procedure following this criterion.

"Systematic-training" condition

In the "systematic-training" condition, a different pair of Re and NoRe males (chosen at random) was assigned to 5 female subjects and the same assignment was then used for all propagation distances tested: a given female was always tested with the same pair of males. In addition, each female was successively challenged as follows: Day 1 = distance calls propagated at 2 m; Day 2 = 64 m; Day 3 = 128 m; Day 4 = 256 m. We hypothesized that this cumulative training from short to medium and long distances would help females to increase the active space of the male signals, in other words, to discriminate between the males in spite of increasing sound degradation.

"No-training" control condition

In the "no-training" condition, we randomly selected four pairs of males from our database and used them as stimuli for all subjects (*n* = 7 females; the same 5 as in the "systematic-training" condition and two additional ones). For a given female, the pair of males and the propagation distance (2, 64, 128 and 256 m) used were randomly assigned across the 4 testing days. Thus, the subjects were all tested with the same males and the same propagation distances, but not in the same order, both parameters being balanced across subjects. Thus, all subjects had to learn to discriminate between a different pair of males every day and therefore had no cumulative training for one set of stimuli from one day to the other. This test provided insight into the baseline capacity for discrimination of degraded calls.

Data analysis

For both protocols, we retrieved for each pecking event the stimulus type (Re/NoRe) and the subject's response (Interruption/Non-interruption). The interruption behaviour of the subject was used as the dependent response variable. Using logistic regression, we tested the effects of the stimulus type, the distance and/or the spectral correlation (i.e. the acoustic similarity) between the Re and NoRe sounds. Subject identity was used as a random factor to take into account potentially different biases in average interruptions across conditions for each bird. Spectral correlations between sound stimuli were calculated using the meanspec function of Seewave R package (Sueur et al., 2008). As described in the results, we found that all three factors (stimulus type, distance, stimulus correlation) were significant. Then, to both visualize the results and to serve as post-hoc tests, we analysed the data for each distance separately and without taking the correlations into account. For each distance and each stimulus type we calculated the probability of interrupting the stimulus averaged across birds. This average probability was obtained from the total number of pecks and the total number of interruptions across birds. Note that this average probability gives higher weights to the birds that pecked more. This is appropriate since our confidence for the probability of interruption for birds that pecked more is higher; also, very similar results were found by first estimating the probability for each bird and then performing the average. Statistical significance was obtained from the logistic regression that predicted interruption probability and used stimulus type as the regressor and bird as a random factor. Using bird as a random factor also allowed us to exclude outlier effects (where one bird would dominate the data). This was not the case in our data since performing the logistic regression without the random effect yielded identical statistical results. If the random factor is excluded, the logistic regression is identical to an exact binomial test for proportion. In figure 2, the statistical significance (shown as *) is obtained from the linear regression with the random effect and the confidence intervals for the probabilities are shown for the exact binomial test.

Acknowledgements

We warmly thank Colette Bouchut and Nicolas Boyer for technical support in the ENES lab, Yuka Minton and Michelle Carney in the Theunissen lab, and Clémentine Vignal for discussion. We are grateful to Alexandra Hernandez for technical and programming help with the Preference test setup.

Funding

This study was funded by the Agence Nationale de la Recherche (A.N.R. project “Acoustic Partnership” to J.E.E., N.M. and S.C.M.), the France-Berkeley Fund (to N.M. and F.E.T.), the National Institutes of Health [grant no. R01DC007293 to F.E.T.], the Fyssen Foundation (to J.E.E.), the French Ministry of Research (PhD stipend to S.C.M.), as well as a Monahan fellowship and a Fulbright fellowship to S.C.M.

Author contributions

All authors participated in the conception and design of the experiments. S.C.M. and J.E.E set up the experiments and collected the data, and S.C.M. and F.E.T. analyzed it. S.C.M. prepared the manuscript with advice and interpretation from all authors.

References

Aubin, T. and Jouventin, P. (1998). Cocktail–party effect in king penguin colonies. *Proc. R. Soc. Lond. B Biol. Sci.* **265**, 1665–1673.

Aubin, T. and Jouventin, P. (2002). How to vocally identify kin in a crowd: The penguin model. In *Advances in the Study of Behavior*, pp. 243–277. Elsevier.

Bee, M. A. (2008). Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. *Anim. Behav.* **75**, 1781–1791.

Blumenrath, S. H., Dabelsteen, T. and Pedersen, S. B. (2007). Vocal neighbour–mate discrimination in female great tits despite high song similarity. *Anim. Behav.* **73**, 789–796.

Brémond, J.-C. (1978). Acoustic Competition Between the Song of the Wren (*Troglodytes Troglodytes*) and the Songs of Other Species. *Behaviour* **65**, 89–97.

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

Brumm, H. (2004). The impact of environmental noise on song amplitude in a territorial bird. *J. Anim. Ecol.* **73**, 434–440.

Butterfield, P. A. (1970). The pair bond in the zebra finch. In *Social behaviour in birds and mammals: Essays on the social ethology of animals and man*, pp. 249–278. London: J. H. Crook.

Dabelsteen, T. and Mathevon, N. (2002). Why do songbirds sing intensively at dawn? A test of the acoustic transmission hypothesis. *Acta Ethologica* **4**, 65–72.

Dent, M. L., McClaine, E. M., Best, V., Ozmeral, E., Narayan, R., Gallun, F. J., Sen, K. and Shinn-Cunningham, B. G. (2009). Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *J. Comp. Psychol.* **123**, 357–367.

Emlen, S. T. and Oring, L. W. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science* **197**, 215–223.

Falls, J. B. (1982). Individual recognition by sound in birds. In *Acoustic Communication in Birds* (ed. Kroodsma, D. E. and Miller, E. H.), pp. 237–278. New York.

Forrest, T. G. (1994). From sender to receiver: propagation and environmental effects on acoustic signals. *Integr. Comp. Biol.* **34**, 644–654.

Francis, C. D., Ortega, C. P. and Cruz, A. (2010). Vocal frequency change reflects different responses to anthropogenic noise in two suboscine tyrant flycatchers. *Proc. R. Soc. B Biol. Sci.* **278**, 2025–2031.

Hulse, S. H. (2002). Auditory scene analysis in animal communication. In *Advances in the Study of Behavior*, pp. 163–200. Elsevier.

Jouventin, P., Aubin, T. and Lengagne, T. (1999). Finding a parent in a king penguin colony: the acoustic system of individual recognition. *Anim. Behav.* **57**, 1175–1183.

Klump, G. M. (1996). Bird communication in the noisy world. In *Ecology and Evolution of Acoustic Communication in Birds* (ed. Kroodsma, D. E. and Miller, E. H.), pp. 321–338. Ithaca.

Kondo, N. and Watanabe, S. (2009). Contact calls: information and social function. *Jpn. Psychol. Res.* **51**, 197–208.

Lohr, B., Wright, T. F. and Dooling, R. J. (2003). Detection and discrimination of natural calls in masking noise by birds: estimating the active space of a signal. *Anim. Behav.* **65**, 763–777.

Maddox, R. K., Billimoria, C. P., Perrone, B. P., Shinn-Cunningham, B. G. and Sen, K. (2012). Competing Sound Sources Reveal Spatial Effects in Cortical Processing. *PLoS Biol.* **10**, e1001319.

Marler, P. and Slabbekoorn, H. (2004). *Nature's Music, the Science of Birdsong*. Elsevier Academic Press. London.

Mathevon, N., Aubin, T. and Dabelsteen, T. (1996). Song Degradation during Propagation: Importance of Song Post for the Wren *Troglodytes troglodytes*. *Ethology* **102**, 397–412.

Mathevon, N., Aubin, T., Vielliard, J., da Silva, M.-L., Sèbe, F. and Boscolo, D. (2008). Singing in the rain forest: how a tropical bird song transfers information. *PLoS ONE* **3**, e1580.

Miller, D. B. (1979). The acoustic basis of mate recognition by female Zebra finches (*Taeniopygia guttata*). *Anim. Behav.* **27**, 376–380.

Mockford, E. J. and Marshall, R. C. (2009). Effects of urban noise on song and response behaviour in great tits. *Proc. R. Soc. B Biol. Sci.* **276**, 2979–2985.

Nemeth, E., Pieretti, N., Zollinger, S. A., Geberzahn, N., Partecke, J., Miranda, A. C. and Brumm, H. (2013). Bird song and anthropogenic noise: vocal constraints may explain why birds sing higher-frequency songs in cities. *Proc. R. Soc. B Biol. Sci.* **280**, 20122798–20122798.

Okanoya, K. and Dooling, R. J. (1987). Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. *J. Comp. Psychol.* **101**, 7–15.

Park, T. J. and Dooling, R. J. (1986). Perception of degraded vocalizations by budgerigars (*Melopsittacus undulatus*). *Anim. Learn. Behav.* **14**, 359–364.

Park, T. J. and Dooling, R. J. (1991). Sound localization in small birds: Absolute localization in azimuth. *J. Comp. Psychol.* **105**, 125–133.

Phillmore, L. S., Sturdy, C. B., Ramsay, S. M. and Weisman, R. G. (1998). Discrimination of auditory distance cues by black-capped chickadees (*Poecile atricapillus*) and zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **112**, 282–291.

Phillmore, L. S., Sturdy, C. B., Turyk, M. R. M. and Weisman, R. G. (2002). Discrimination of individual vocalizations by black-capped chickadees (*Poecile atricapilla*). *Learn. Behav.* **30**, 43–52.

Phillmore, L. S., Sturdy, C. B. and Weisman, R. G. (2003). Does reduced social contact affect discrimination of distance cues and individual vocalizations? *Anim. Behav.* **65**, 911–922.

Pohl, N. U., Leadbeater, E., Slabbekoorn, H., Klump, G. M. and Langemann, U. (2012). Great tits in urban noise benefit from high frequencies in song detection and discrimination. *Anim. Behav.* **83**, 711–721.

Radziwon, K. E., Welch, T. E., Cone, J. P. and Dent, M. L. (2011). Identification of auditory distance cues by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *J. Acoust. Soc. Am.* **129**, 3384–3392.

Slabbekoorn, H. (2013). Songs of the city: noise-dependent spectral plasticity in the acoustic phenotype of urban birds. *Anim. Behav.* **85**, 1089–1099.

Slabbekoorn, H. and Peet, M. (2003). Ecology: Birds sing at a higher pitch in urban noise. *Nature* **424**, 267–267.

Sueur, J., Aubin, T. and Simonis, C. (2008). Equipment review-Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics* **18**, 213–226.

Vignal, C., Mathevon, N. and Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature* **430**, 448–451.

Vignal, C., Mathevon, N. and Mottin, S. (2008). Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behav. Processes* **77**, 191–198.

Warren, P. S., Katti, M., Ermann, M. and Brazel, A. (2006). Urban bioacoustics: it's not just noise. *Anim. Behav.* **71**, 491–502.

Wiley, R. H. and Richards, D. G. (1982). Adaptations for acoustic communication in birds: sound transmission and signal detection. In *Acoustic Communication in Birds* (ed. Kroodsma, D. E. and Miller, E. H.), pp. 131–181. New York.

Zann, R. (1984). Structural variation in the zebra finch distance call. *Z. Für Tierpsychol.* **66**, 328–345.

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

Auditory scene analysis by the songbird brain: Specialized neurons are expert at discriminating individual vocal signatures despite extreme acoustic constraints

Solveig C. Mouterde[1], Julie E. Elie[2], Nicolas Mathevon[1] & Frédéric E. Theunissen[2]

[1] Equipe de Neuro-Ethologie Sensorielle ENES-CNPS CNRS UMR8195, Université de Saint-Etienne, 42023 Saint-Etienne, France

[2] Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720

Original manuscript

Question

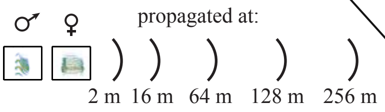
Is there a neural substrate for discrimination of degraded voices in the zebra finch's brain?

Methods

Electrophysiological recordings of neural responses to degraded calls

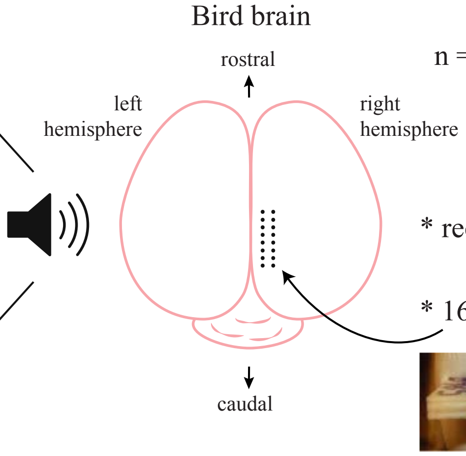
Stimuli

distance calls of
unfamiliar birds



Con stimuli: naturally propagated calls

Syn stimuli: same intensity decrease as Con stimuli but constant SNR



Subjects

n = 8 male and female adult birds

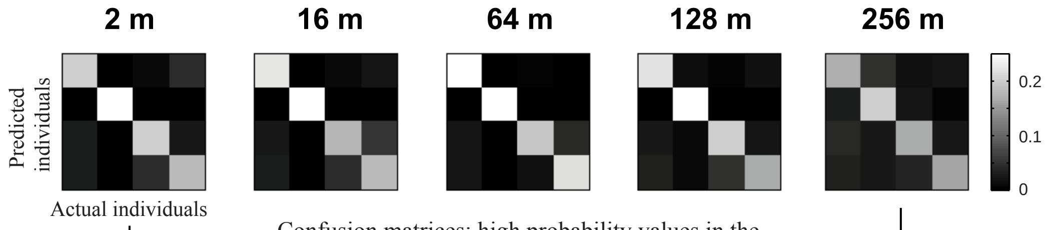
Protocol

* recordings on anesthetized birds

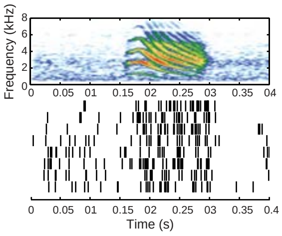
* 16-channel multi-electrode array

Results

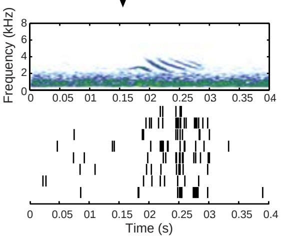
We found units that discriminate individual vocal identities in degraded calls, without prior familiarization of training:



Confusion matrices: high probability values in the diagonal shows high discrimination performance



Spike trains for the 8 trial presentations of the same call



* these units are non selective:
they discriminate as well between all individuals

* they do as well for Con and Syn calls:
invariance to the reduction of the signal-to-noise ratio

* they use the same code for each individual across distances:
categorization of individual identity regardless of distance

Auditory scene analysis by the songbird brain:
Specialized neurons are expert at discriminating individual vocal signatures despite extreme acoustic constraints

Solveig C. Mouterde, Julie E. Elie, Nicolas Mathevon and Frédéric E. Theunissen

Abstract

One of the most complex tasks for the brain is to categorize the vast amount of stimuli that it continually receives into behaviorally relevant objects. Although "scene analysis" is a universal problem solved by animals, studying it computationally and experimentally has been a challenge for scientists. The challenge is particularly demanding in *auditory* scene analysis where spatial effects such as location and distance are intertwined with source effects such as multiple speakers and noise as all propagated sounds sum to form a single acoustical signal. To address the full complexity of the task, researchers need to rely on ecologically relevant stimuli and behaviors. Using a songbird model, the zebra finch (*Taeniopygia guttata*), for which finding its partner in constraining communication conditions is an everyday issue in the wild, we investigated the neural substrate for discrimination of individual vocal signatures when signals have been degraded through propagation in a natural environment.

Performing electrophysiological recordings on adult anesthetized birds, we found a population of neurons in the avian auditory forebrain that discriminate individual voices with various degrees of propagation-induced degradation. These highly discriminant neurons showed low selectivity towards different individuals, and could categorize individual identity regardless of propagation distance. We also showed that, after taking into account intensity effects, the highly discriminant neurons were not affected by additional propagation-induced degradations. This study shows for the first time the presence of high-level distance invariant neurons in the primary auditory cortex: neurons that are able to discriminate the vocal signature irrespective of propagation distance. These neurons having been recorded in naive animals that had no prior experience with the voices tested, we suggest that this highly discriminant population of neurons as part of a general purpose system for voice discrimination, that could be enhanced by perceptual learning.

INTRODUCTION

One of the biggest tasks for the brain is to discriminate, in the midst of the prodigious amount of stimuli that it continually receives, what is relevant from what is not. This task is further constrained by the multiple sources of noise in natural environments that contribute to the degradation of the biologically relevant information. Although real world "scene analysis" is a universal problem solved by all animals (Appeltants et al., 2005; Aubin and Jouventin, 2002; Hill, 1979; Schnitzler and Flieger, 1983; von der Emde, 2004), scientific research has had tremendous difficulties in dealing with the complexity of the process (Bregman, 1993). This issue has been approached by investigating the neural basis of perceptual invariance (i.e. the stability in the perception of stimuli despite large variations in their parameters; Billimoria et al., 2008), yielding significant insight into the fields of visual (Logothetis and Sheinberg, 1996; Quiroga et al., 2005), auditory (Billimoria et al., 2008; Moore et al., 2013; Schneider and Woolley, 2013) and olfactory perception (Stopfer et al., 2003). However, a major limitation in current neurophysiological approaches is that they do not acknowledge or address the complexity of the problems that need to be solved (Lewicki et al., 2014): experiments based on simplistic models of animal behavior, or using "idealized" stimuli that often do not reflect ecologically relevant stimuli and behaviors, have little chance of providing a thorough understanding of the mechanisms at play in more complex natural settings. In this respect, one overlooked aspect in auditory scene analysis concerns how the brain extracts information required for social interactions when information transfer between individuals is under high external constraints. Specifically, although individual recognition is essential for many animals' social life, the neural substrate allowing discrimination between individual voices when the emitter and the listener are remote and the communication signals are affected by propagation is not known. Here we address this question by investigating the real-life issue of partner acoustic recognition in a songbird where paired mates commonly experience extreme sound degradations induced by long-range propagation.

The zebra finch (*Taeniopygia guttata*), a model system showing striking similarities to humans in the context of speech (Doupe and Kuhl, 1999), is a small gregarious songbird from subarid Australia that forms pair bonds for life and lives in large flocks in open country with a scattering of trees and bushes (Butterfield, 1970; Zann, 1996). Because these birds are opportunistic breeders living in a very unpredictable environment (Zann, 1996), maintaining strong pair bonds between breeding events is of utmost importance. With the flock being constantly on the move, using a solid vocal recognition system allows partners to avoid the cost of losing each other. Zebra finches thus offer a unique opportunity for studying how the male and female emitter's vocal identity is degraded through propagation in a natural environment, and how these birds extract information from these degraded signals at the neural level.

Among the vocal repertoire of zebra finches, the distance call –a short call emitted by both males and females– is the ideal candidate for our research, as it is used especially within mate-bonded pairs while foraging and when a visual connection has been lost (Zann, 1996). Previous studies have shown that the distance call bears an individual signature, and that the birds are capable of call-based individual recognition (Vignal et al., 2004; Vignal et al., 2008; Zann, 1984). As calls propagate through the environment, their quality is degraded, affecting the signal's amplitude as well as its spectral and temporal structure (Wiley and Richards, 1982; Mouterde et al., *submitted*). However, it has been experimentally demonstrated that female zebra finches can discriminate between the calls of males propagated at 128 m, and that this performance can be extended to more than 250 m if they have the opportunity to train themselves with these stimuli (Mouterde et al., *in revision*). This impressive perceptual ability must be subserved by auditory neurons whose responses are only minimally affected by propagation-induced sound degradation.

The neural representation of propagation-induced degraded vocalizations has still to be explored. To this end, we chose to focus on individual vocal signature as it requires particularly fine spectral and temporal information (Charrier et al., 2003; Latinus et al., 2013; Mathevon et al., 2008), which can help us understand further the subtleties of neural processing. In a previous study (Mouterde et al., *submitted*), we showed that individual identity in distance calls of zebra finches is encoded redundantly, with some parameters being more robust to propagation-induced degradation than others, and the most informative parameters being different at short and long distance. The extraction of individual identity in degraded vocalizations should therefore require the use of some the most refined tuning properties of the auditory system; evidence for the ability of auditory neurons to represent subtle spectro-temporal features in songbirds has been shown by the extensive research in this area (Sen et al., 2001; Theunissen and Shaevitz, 2006; Theunissen et al., 2000; Woolley et al., 2005). Furthermore, research on auditory scene analysis had led to investigate specialized tuning for most robust features, and thus to study invariance properties of neurons. Most published research to date has only indirectly examined this question by analyzing how the neural representation of song varies as a function of intensity or is affected by noise (Billimoria et al., 2008; Boumans et al., 2008; Moore et al., 2013; Narayan et al., 2007). Compared to calls, songs contain an added layer of information, the syntax (i.e., the arrangement of syllables typical of each male or male lineage), which can also be used in songbirds to perform individual discrimination (Gentner, 2004). Disentangling to which extent the neural processing uses the fine spectro-temporal features of the signal or its syntactic organization to characterize an individual's signature may therefore be challenging when using songs as stimuli. Furthermore, calls are far better candidates than songs to study propagation-induced degradation since distance calls are used for long distance communication while the directed song is only used in proximity to a female in zebra finches.

Additionally, they support voice-based recognition between individuals of both sexes in natural context. Studies using songs however yielded important cues about which brain areas are possibly involved in vocal discrimination, and indicate an increased tolerance for noise as one ascends the auditory pathway. FMRI-based evidence suggested that the first region within the songbird auditory system where the ability to discern a song from masking noise emerges is located in the caudomedial nidopallium (NCM), a secondary auditory area (Boumans et al., 2008). This finding was corroborated by two electrophysiology-based studies on zebra finches, the first one studying noise-invariant responses in the NCM using unfamiliar conspecific songs (Moore et al., 2013), and the second describing a population of neurons in the same region that encoded previously learned individual songs embedded in background chorus noise (Schneider and Woolley, 2013). Another study described a population of neurons from the thalamo-recipient field L (a region that is analogous to the primary auditory cortex of mammals) that are able to discriminate between bird songs while being invariant to intensity: the neurons with a high degree of invariance also displayed a high discrimination performance (Billimoria et al., 2008). This neural property could also play a key role for the task of discriminating between calls of different individuals emitted at varying distances in a natural environment.

Other research in songbirds has shown high selectivity for conspecific songs in the auditory forebrain, in the field L (Grace et al., 2003; Hauber et al., 2007; Theunissen and Shaevitz, 2006), as well as in the caudal mesopallium, CM (Grace et al., 2003; Hsu et al., 2004b) and the NCM (Bailey et al., 2002; Ribeiro et al., 1998). These studies and others suggested that hierarchical sensory processing occurs in the auditory cortex, with the secondary auditory areas (CM and NCM) showing sensitivity to higher-order features: neural substrates for individual recognition of songs in the auditory cortex were found in the medial CM (CMM; Gentner and Margoliash, 2003) and the NCM (Chew et al., 1996; Gentner, 2004). One neurophysiological study using calls instead of songs demonstrated that calls from individuals differing in their degree of familiarity with the tested bird evoked differential neuronal responses in the NCM (Menardy et al., 2012). Overall however, our knowledge about the brain structures and neural processing involved in voice-based individual recognition remains scarce. And specifically, for the aims of this study, nothing is known on how the songbird brain deals with the impact of sound degradation during individual vocal discrimination.

In the present study we quantified the neural discriminability of single neurons in the zebra finch auditory forebrain for individual vocal signature in the distance call, and investigated how discrimination performance was affected by degradation from propagation with stimuli recorded in natural conditions. For this purpose, we addressed the following questions: does discrimination between individual vocal signatures in degraded signals exist at the neuronal level, without prior familiarization or learning? If such discriminant neurons exist, are they localized in specific areas of the auditory cortex? Do their responses to the calls after a long-range propagation through a natural environment differ from the responses to a mere sound intensity decrease? And finally, is there an effect of sex on the discrimination performance, either at the level of the stimulus (male or female call) or at the level of the perceiver? To investigate these questions, we performed extracellular electrophysiological recordings on zebra finches, using multi-channeled electrode arrays, and obtained single unit responses to distance calls with different levels of propagation-induced degradation.

MATERIALS AND METHODS

Stimulus design and recordings

The stimuli included natural and synthetic calls of male and female zebra finches. The natural calls were distance calls from unrelated and unfamiliar conspecifics (Con) that had been propagated at various distances (from 2 to 256 m) and recorded in natural conditions. In order to disentangle the sheer effect of signal intensity decrease from the effect of the decrease of the signal-to-noise ratio on neural responses (SNR; here defined in the broad sense, i.e. taking into account both the effects of the decrease of the signal's amplitude comparatively to the relatively constant amplitude of the background noise as well as the degradation of the call's temporal and spectral structure due to propagation though the environment), the subjects were also tested with synthetic stimuli (Syn) for which the sound intensity matched that of the Con calls at each propagation distance, but the SNR was constant (see details below).

To prepare the Con stimuli we used a database of distance calls recorded from 32 zebra finches (16 females, 16 males) raised in the ENES laboratory. The calls were recorded in a soundproof room using a microphone (Sennheiser MD-42, Wedemark, Germany) placed 0.2 m above the cage and connected to a Marantz Professional Solid state recorder (PMD 670, Eindhoven, The Netherlands; sampling frequency: 44100 Hz). Each bird was recorded in the presence of two females placed 3 m away and used as an audience to minimize stress; the bird was stimulated with distance calls playbacks from previously recorded conspecific birds. These experimental protocols were approved by the Jean Monnet University's animal care committee (authorization n°42-218-0901-38 SV 09 to ENES lab). We recorded 16 different calls from each individual to make our calls database (total number of calls: 16 * 32 = 512 calls). All the calls were normalized to each other by matching the maximum values of the sound pressure waveforms. We recorded the propagated sounds from all the calls of this database in natural conditions on an open flat field in France. For these sound recordings, the calls were broadcast from a portable solid state recorder (Marantz PMD671) connected to a MegaVox speaker (PB-35W) placed on a stool, 1.30 m high. The speaker volume was set to obtain a sound level of 70 dB SPL at 1 m (Vignal et al., 2008). The sounds were recorded with a Schoeps microphone (MK4 cardioid, on a CMC6-U base) equipped with a Schoeps Basket-type Windscreen (W 20) and set 1.30 m high. The microphone was connected to a second Marantz recorder (PMD671; sampling frequency: 44100 Hz). The propagated calls were recorded 2 m, 16 m, 64 m, 128 m and 256 m away from the source. From these recordings, we isolated 16 different calls per individual per propagation distance (16 calls * 32 individuals * 5 distances = 2560 calls). For the neural recordings, the propagated calls were high-pass filtered with a cut-off frequency of 500 Hz. This frequency cut-off is below the lower frequency threshold of the zebra finch's audiogram (Okanoya and Dooling, 1987) and this filtering was used to remove irrelevant background noise.

To create the Syn stimuli, we used the set of Con calls recorded at 2 m and reduced the gain of each call so that its amplitude matched the amplitude of the same call that had been propagated at each other distance. To do this, we centered the calls inside a time window of the length of the longest call, for males and females separately, then used on these centered calls another time window corresponding to the length of the shortest call for each sex to match the amplitude of the Con call at 2 m to the amplitude of the corresponding call propagated at other distances. In this way we ensured that this amplitude matching always involved sounds in which the signal was present, regardless of the varying durations of calls from different individuals. We thus obtained Syn calls for 4 distances (16, 64, 128 and 256 m) for which the SNR was the same as the 2 m call but the amplitude matched the amplitude of the Con call at these 4 distances. This database consisted of 16 different calls per individual for the last 4 propagation distances (16 calls * 32 individuals * 4 distances = 2048 calls).

Animal procedures and electrophysiological recording protocol

Eight adult zebra finches (n = 4 males and 4 females) were used in the electrophysiology experiments. Extracellular recordings from ensemble of single units were obtained in the

Urethane anesthetized subjects, immobilized on a stereotaxic apparatus, using one or two 16 channel array electrodes (Omn1010, Tucker-David Technologies Inc., Alachua, FL, USA) consisting of 2 rows of 8 electrodes (width 0.5 mm, length 2 mm, distance between two electrodes within one row: 250 µm). The electrode arrays were lowered into the auditory forebrain using a microdrive. Each electrode had previously been coated with DiI stain (Invitrogen, Eugene, Oregon, USA) so as to facilitate the electrodes localization in the brain during the histological analysis. The recording took place in a double-walled anechoic chamber (Acoustic Systems, MSR West, Louisville, CO, USA) where a loudspeaker (Blaupunkt T-line, Berlin, Germany) was used to broadcast the stimuli. The volume of the loudspeaker was set to deliver zebra finch calls at 70 dB SPL (Digital Sound Level Meter, RadioShack, weighting type B) and was placed 20 cm in front of the subject's head. Extracellular voltages were recorded with a system from Tucker-Davis Technologies (TDT). All animal procedures were approved by the Animal Care and Use Committee at University of California Berkeley.

We recorded multiunit responses from each electrode in the array at two to six recording depths per subject. These multiunit recordings were then spike sorted off-line as described below. Out of the 8 total subjects, 6 were tested with a single array in one hemisphere (4 on the right, 2 on the left), and 2 were tested simultaneously with two arrays, one in each hemisphere. The electrode arrays spanned the mediolateral (from 0.25 to 1.5 mm lateral from the y-sinus) and the rostrocaudal (from 0.25 to 2.1 mm rostral from the y-sinus) axes of the auditory forebrain. The depth of the recording sites spanned between 1.15 and 2.13 mm from the brain surface, and the minimum distance between two sites was 100 µm. For each electrode in the array, spike arrival times and spike waveforms (snippets) were recorded by thresholding the extracellular voltage trace.

For each site, calls from 4 or 8 different individuals from the same sex were broadcasted to the subject; sex was swapped from one site to another. The call identities selected for each site (i.e., individual name and call number) were selected randomly from the database, and for each call identity, the calls recorded at all 5 distances were used for the Con stimuli, and (depending on the protocol type -see below) the synthetic calls for all 4 distances were used for the Syn stimuli. In order to limit the recording time for each site, the number of calls selected per individual varied for each condition: two subjects were only tested with Con stimuli ('Con-only' protocol) and heard 8 different calls from 8 individuals at each site (that is, a total number of 8 individuals * 8 calls * 5 Con distances = 320 different calls per site); for the other six subjects, tested with both stimuli ('Con+Syn' protocol), those tested with 4 individuals heard 8 different calls per individual, and those tested with 8 individuals heard 4 different calls per individual (that is, a total number of

4 or 8 individuals * 8 or 4 calls * 5 Con distances + 4 or 8 individuals * 8 or 4 calls * 4 Syn distances = 288 different calls per site). The calls had a mean duration of 0.22 s ± 0.06 s for females and 0.13 s ± 0.05 s for males, and a mean fundamental frequency of 549.7 Hz ± 126.5 Hz for females and 803.3 Hz ± 164 Hz for males. The inter-stimulus interval was uniformly distributed between 1 to 3 s to prevent any rhythmic pattern that could potentially entrain the neurons or generate expectations. We repeated the presentation of all the stimuli 8 times and for each of these 8 trials the order of stimulus presentation was randomized. In this manner, we avoided any stimulus dependent adaption.

Histology and anatomical localization of electrodes

After the recording, the bird was euthanized with an overdose of isoflurane and transcardially perfused with phosphate buffered saline (PBS, Sigma Chemical co., St Louis, MO, USA), followed by 4% formaldehyde. The brain was sunk in 4% formaldehyde followed by 30% sucrose, before being frozen using liquid nitrogen in preparation of the histological procedures. The brain was then sliced frontally or parasagitally in 20-μm-thick sections using a freezing microtome. Alternating brain sections were stained with either cresyl violet or DAPI nucleic acid stain, to be used to visualize electrode tracks (with the DiI stain marking each electrode emplacement) and histological tissues. These observations were made using a Zeiss AxioImager M2 fluoresence microscope fitted with a camera (Retiga 1350 EX, QImaging).

Localization of the electrodes involved measuring the distance from the entry of the electrodes to their deepest point and comparing it to the depth of the last recording site as shown on the microdrive used during the recording; recording site localization could then be achieved from the coordinates of each site obtained from the microdrive. Using well-known landmarks such as the *lamina mesopallialis* (LaM; known in the old nomenclature as hyperstrial lamina or LH) and the dorsal *lamina pallio-subpallialis* (LPS; previously called the medullary lamina or LMD) and differences in cell density as described in literature (Fortune and Margoliash, 1992), recording sites where then assigned to either CM (lateral: CLM, or medial: CMM), NCM, thalamo-recipient subdivision L2b or subregions L1 or L of the field L complex. Further distinctions in the field L complex (L2a and L3) were not achieved with enough certainty to be assigned as such, and recording sites found in undefined areas were labeled Lx. Following guidelines found in literature (Vates et al., 1996), we chose to approximate the limit between CMM and CLM as being 800 μm away from the midline.

Data processing and analysis

Pre-processing of the data

All multiunit recordings were sorted into single unit data based on the spike shapes. This spike sorting was performed using a semi-automatic custom made program written in Matlab (Mathworks) that used both un-supervised (k-means) and supervised clustering algorithms (Random Forest). In a first stage, templates for single spike shapes were chosen by the user using a GUI and exploratory cluster analysis. For this process, random groups of 200 spikes were clustered into 6 groups using k-means algorithm. The k-means clustering used the coefficients of a Principal Component Analysis (PCA) on the spikes shapes. The user could then manually select the groups of spikes that corresponded to isolated samples of the same single-unit and assign these as templates for units to be sorted. The process was repeated to find new units or to add more clean samples to particular single-units templates already chosen. In a second stage, these templates were used to train a Random Forest (RF) that used the PCA coefficients and additional spike parameters: the max and min amplitude and the peak slope. The trained RF was then used to classify the remainder of the spikes into single units, noise or non-classifiable units (multi-unit). From 640 multi units recordings, we extracted 2467 single units or an average of 3.85 units per site and electrode. All subsequent analyses were done on this data set of single units.

A unit was defined as auditory (i.e., responsive to acoustic stimuli) if it had reproducible spike patterns in response to the same stimuli. To quantify "reproducible spike patterns" we estimated the coherence between the average response from one half of the trials to the average response of the other half of the trials (Hsu et al., 2004a). This coherence was bias corrected and standard error estimates were obtained using jackknife procedure and multi-tapered estimates of the cross and self spectra (Efron, 1982). Furthermore, we used similar calculations on neural responses to silence to obtain an upper bound on values that could obtained by chance even after bias correction. Non-auditory units (n = 52) were excluded from further analysis.

Characterization of the neural responses

From the auditory single units, we calculated for each recording site the mean firing rate during stimulus exposure, as well as a z-score value for the spike rate, characterizing the normalized difference between the stimulus-evoked mean firing rate and that of the background activity preceding the stimulus.

In addition to the firing rate description, we quantified the neural discriminability of single units for calls from individual birds using an optimal decoder procedure using complete spike patterns (Amin et al., 2013; Gastpar et al., 2010; Wang et al., 2007). In brief, spike templates were obtained for each stimulus using 7 random selections out of the 8 trials, to calculate an average response. The remaining spike trial for each stimulus was then compared to all other trials by calculating the Euclidian distance between the spike test trial and all the average templates (van Rossum distance; Rossum, 2001). Before calculating this distance both the test trial and the templates were convolved with a decaying exponential with varying time-constant. By using very large time-constants, neural discrimination would be based solely on the mean firing rate. Very short-time constants would capture fine temporal structure in spike patterns that potentially code stimulus information but that could also be neural variability. The stimulus template yielding the smallest distance designated the stimulus that was decoded. This process was repeated over 1000 iterations of random selections to obtain a confusion matrix showing the joint probabilities of the predicted vs. actual responses. For each unit, a confusion matrix was estimated for 7 different time-constants (ranging from 1 to 100 ms) for the Con stimuli at 2 m in order to find the time scale giving the best discrimination; this time scale (calculated separately for each neuron) was applied for all other stimuli (Con from 16 to 256 m and Syn if applicable), which enabled us to maximize the discrimination performance for each unit depending on its neural code (Chicharro et al., 2011). Ultimately we obtained a single confusion matrix that showed the best performance for decoding a group of stimuli for a particular singular unit: a confusion matrix was calculated for the Con and Syn stimuli, and for each propagation distance separately.

Since we were ultimately interested in the neural discrimination of the individual producing the calls, we further collapsed the confusion matrix by grouping all stimuli produced by the same individual. To summarize and quantify the discriminability we used two measures. First, we calculated the mutual information (MI) of this reduced confusion matrix; the MI of a uniformly distributed confusion matrix tends towards zero, and the MI of a highly organized matrix (e.g., if the highest probabilities are found in the diagonal) tends towards its maximum: the $\log_2$ of the number of categories (in this study 2 bits for a matrix comparing 4 individuals and 3 bits for 8 individuals). Second, using only the diagonal the confusion matrix, we obtained the percentage of correct classification (PCC) at each distance.

Selection of units discriminant for individual identity

One of our goals was to assess how neural discrimination for individual identity was affected by propagation. For this purpose, we first selected the units that could perform this task at close distance (i.e. at 2 m). We will refer to these units as the discriminant units, bearing in mind that this label does not inform about the quality of the discrimination, i.e. that it refers to any unit that showed any amount of discriminability above chance. This chance level was calculated using the bootstrapping resampling technique: for each neuron, we obtained a bootstrap version of its confusion matrix at 2 m by randomizing the assignment of the individuals to the calls. From 1000 of such "identity-scrambled" confusion matrices we obtained a distribution of MI values that could be obtained by chance (MI_{rand}). We then compared the observed MI value (MI_{obs}) to this distribution of MI_{rand} values, following the null hypothesis that MI_{obs} was not different from the distribution, and calculated a z-score value, as follows:

$$z_{MI} = \frac{MI_{Obs} - \mu_{MI_{Rand}}}{\sigma_{MI_{Rand}}}$$

where $\mu_{MI_{Rand}}$ is the mean of the MI_{rand} distribution, and $\sigma_{MI_{Rand}}$ its standard deviation. We then calculated the probability density function p for z_{MI}, and rejected the null hypothesis (i.e. defined the tested unit as discriminant) for $p < 0.01$. Out of the 2415 total number of single auditory units, we retrieved 1532 discriminant units. Non-discriminant units were excluded from further analysis.

Next, in order to compare the MI values obtained from subjects tested with 4 or 8 different individuals (MI_{obs}), and to eliminate any positive bias in the evaluation of discrimination performance, we calculated a corrected value for the MI (Chicharro et al., 2011) as follows:

$$MI_c = MI_{Obs} - \mu_{MI_{Rand}}$$

We performed this bias correction for information calculated at all distances. For non-significant neural discrimination values, this quantity can become negative. These negative values were set to zero but effectively eliminated from the analysis as described below. All the MI values presented in the Results section and figures are corrected values (MI_c).

Investigation of invariance properties to sound degradation

To characterize the effect of propagation distance on neural discrimination, we calculated MI_c for each distance (for Con and Syn calls and for each discriminant unit) and fitted an exponential model on these curves to obtain a measure of "discrimination decay" in meters^{-1}. In order to compare the discrimination decay values for the Con and Syn stimuli, the model fitted on the Syn curves included the Con value at 2 m, so as to have the same number of distance points in both cases (the Syn stimuli being derived from the Con value at 2 m in the first place). For this fit, we only used distances up to the first (and not including) distance where MI_c was zero (or all the distances if $MI_c > 0$). In the case when only one MI_c value was available, no discrimination decay value could be retrieved (this concerned 163 units out of 1532 total for the Con stimuli and 159 units out of 1273 total for the Syn stimuli).

To investigate the neurons' discrimination performance we also calculated a corrected value for the percentage of correct classification (PCC) by subtracting the chance level (25% for subjects tested with 4 individuals and 12.5% for 8 individuals) to each value. This enabled us to compare PCCs from individuals tested with 4 and 8 individuals, and these PCCs over chance level are used throughout the Results section and in all the figures.

Finally, in order to investigate if the neural coding for each individual was the same regardless of distance, for each discriminant unit we calculated a new confusion matrix taking into account the calls propagated at all distances for each individual, and collapsed these "all-distances matrices" by grouping all stimuli produced by the same individual. We then calculated the PCC over chance level for each matrix.

RESULTS

Neural substrate for the spontaneous discrimination between individual voices

Discrimination performance

We found neurons showing high discrimination for individual identity, at short as well as at long distances. As an example, Fig. 1 shows the confusion matrices for a unit that demonstrated a high mean percentage of correct classification (PCC) over the chance level across all five distances.

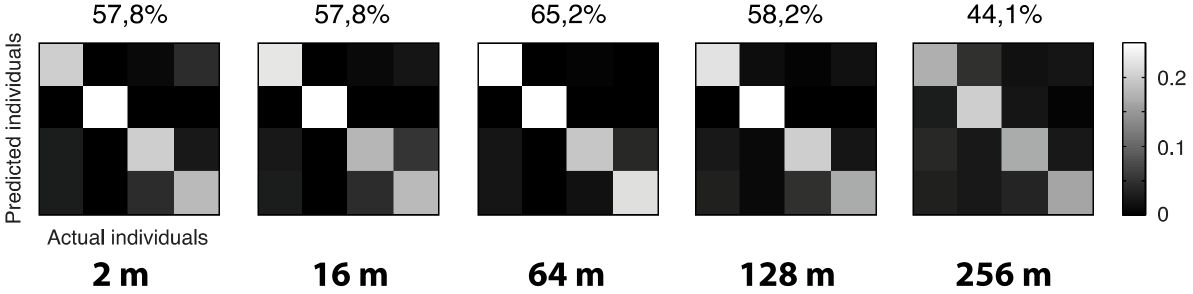


Fig. 1: Confusion matrices for a highly-discriminating unit tested with 4 individuals, at all distances. Probability values in the diagonal indicate a high discrimination ability. The percentage value on top of each matrix is the mean percentage of correct classification over the chance level for this matrix.

Not all neurons showed such discrimination abilities, as illustrated in Fig. 2 where we compare the responses of two units to a stimulus call emitted by the same individual. In this example, for the highly discriminating unit (Fig. 2a) the firing rate during the stimulus is clearly higher than before or after the stimulus at all distances, while for the second unit (Fig. 2b) this difference is visible up to 16 m but not at further distances (from 64 to 256 m). Moreover, the confusion matrices, showing the discrimination ability of each unit for the calls of every individual presented, illustrate the difference in discrimination performance between both units. Next, in order to visualize the discrimination ability of the whole population of neurons, we plotted the discriminant units as a function of the MI_c at 2 m and the discrimination decay (Fig. 3), showing either the total MI across all distances (i.e., the sum of MI_c values for each distance, for each unit) or the mean PCC above chance level across all distances. In this space, the most performing units (i.e., units for which discrimination is high at 2 m and decreases slowly with distance) would have a high MI_c at 2 m and a low discrimination decay (i.e., a slow decrease of MI with distance). Indeed, Fig. 3 shows that units situated close to zero on the y-axis but with high MI values at 2 m do have the highest total MI and mean PCC values. This validates our choice to use either the MI or the PCC to evaluate the neurons' ability for individual discrimination.

Finally, we found that the neural coding for each individual was very similar for all distances for most discriminant units. Fig. 4a shows that the PCC values obtained from the "all-distance matrices" were highly correlated to the mean PCC values calculated separately for each distance. A linear model without intercept on this data confirmed this result (slope = 0.75, t = 169.0, $p < 10^{-4}$, $R^2 = 0.95$). This is further illustrated in Fig. 4b where the "all-distances" matrix and the corresponding collapsed matrix regrouping each individual are shown for a highly discriminant unit: the squares in the diagonal represent each exemplar of the calls used per individual, at each distance, and the fact that the highest probabilities can be found in this diagonal shows that the neural code could predict reliably individual identity in all these call exemplars propagated at different distances.

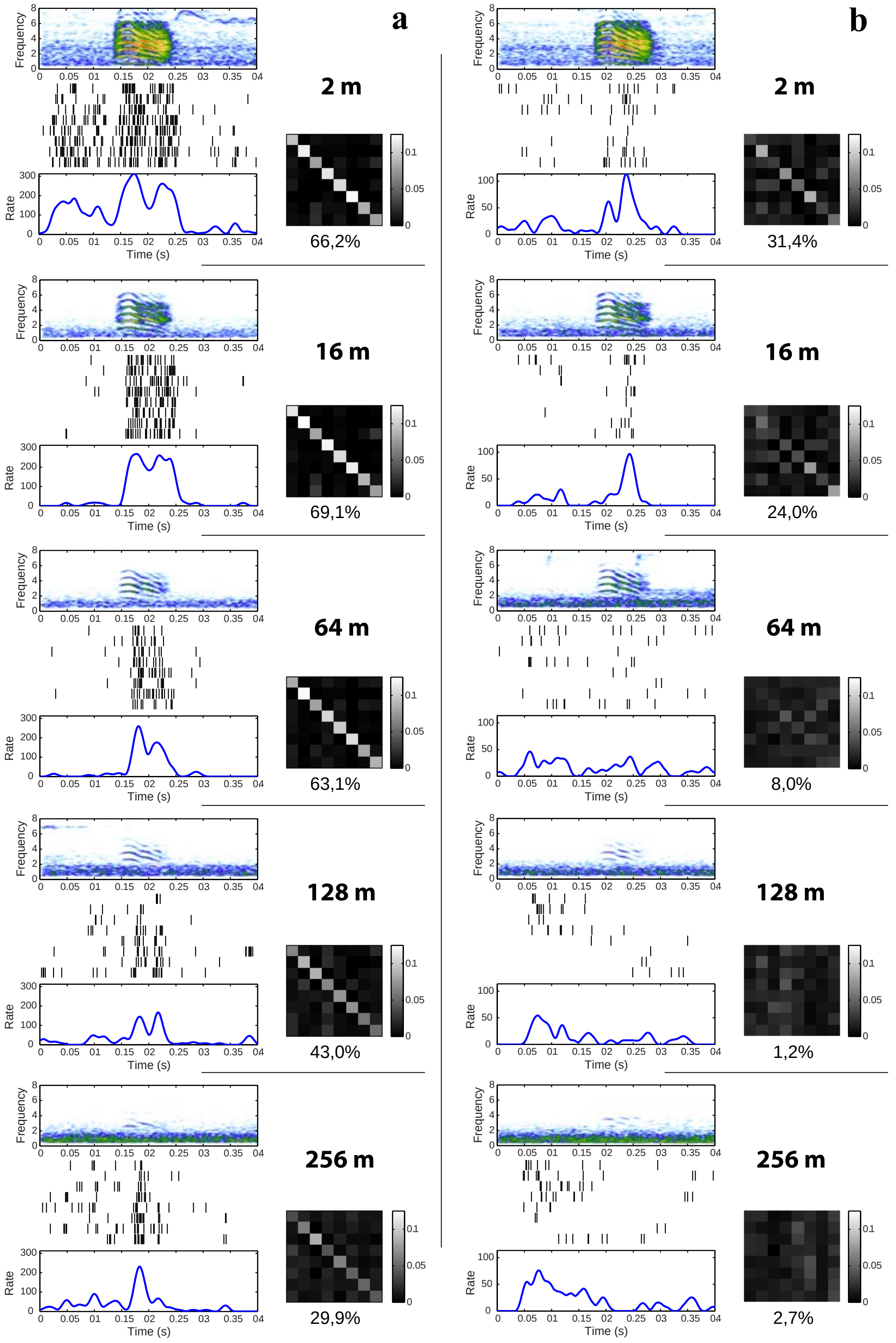


Fig. 2: Comparison of the responses of two units to different male calls after propagation through the environment (Con stimulus). For each propagation distance, the neural responses are shown for a highly-discriminant unit (**a**) and another unit showing discrimination abilities at short and medium distance but not at long distance (**b**). The two units were recorded on different birds. For each distance, the spectrogram of the stimulus call is shown on top, followed by the spike trains for the 8 trials, and by the PSTH averaging these 8 trials. Sound frequency is given in kHz and rate in spikes/s. The confusion matrix shows the discrimination ability between individuals: the whiter the diagonal, the higher the individual discrimination. The percentage value below each matrix is the mean percentage of correct classification over the chance level.

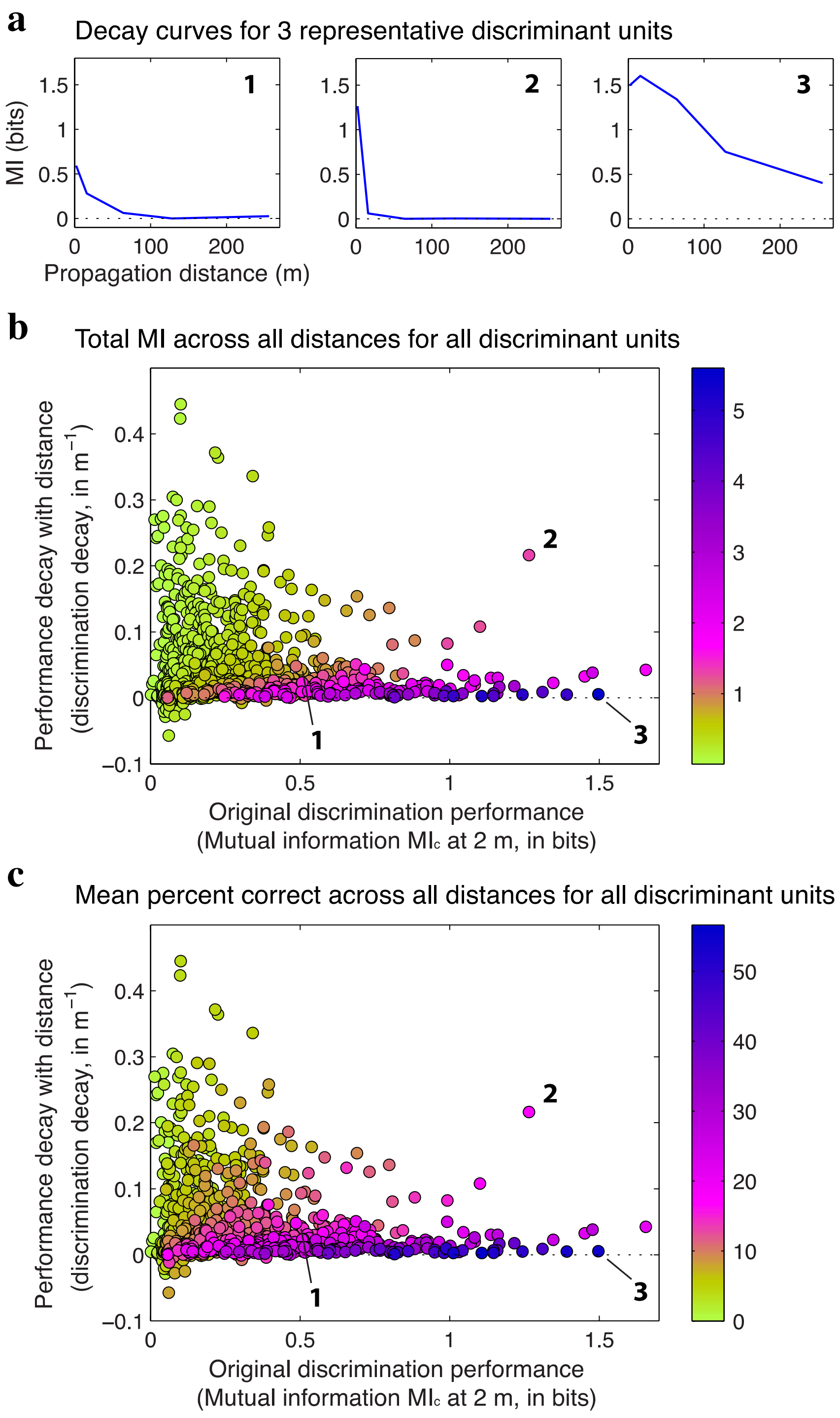


Fig. 3: Between-individuals discrimination performance of the population of discriminant units. The discrimination decay curves (i.e., the discrimination performance MI_c as a function of propagation distance) are shown for three representative units (**a**): with a low MI_c value at 2 m and a medium discrimination decay (1), with a high MI_c value at 2 m and a high decay (2) and with a high MI_c value at 2 m and a low decay ("highly-discriminant" unit, 3). These three units are placed on the two following graphs, in which all discriminant units are plotted as a function of their discrimination performance MI_c at 2 m and discrimination decay (i.e., MI_c decay with distance), with the total MI across all distances (**b**) or the mean percentage of correct classification above chance level across all distances (**c**) represented on the color scale. In this representation, the most discriminant units are situated on the bottom right of the graph, i.e. have high discrimination ability at 2 m and a low discrimination decay with increasing distance. All values are shown using responses to Con stimuli.

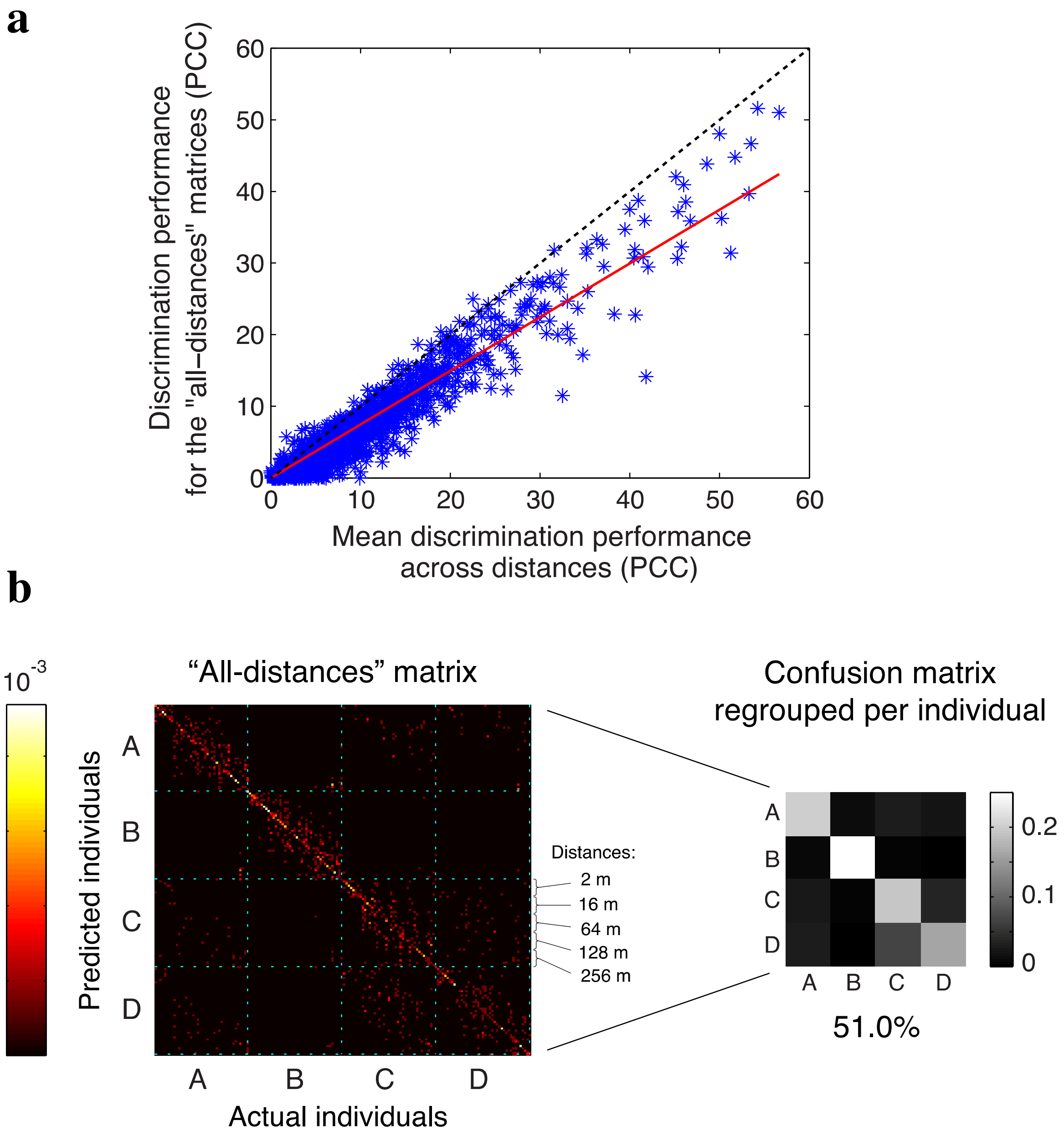


Fig. 4: Coding for individual identity regardless of propagation distance. (a): PCCs calculated from the "all-distances" confusion matrices regrouped by individual as a function of the mean PCCs calculated separately for each distance. The regression line calculated from the linear model without intercept is represented in red. The diagonal x = y is represented as a black dotted line for visual comparison. **(b)**: representation of the "all-distance" matrix and its corresponding collapsed matrix per individual (on which the PCC is calculated) for a highly discriminant unit. Dotted lines on the left matrix represent the squares on the right matrix; on the left matrix, any correct prediction for each individual (A to D) adds value in the squares in the diagonal. For each square (corresponding to one predicted/actual individual assessment), calls are organized sequentially with increasing distances. The percentage value on the right is the PCC over the chance level for the collapsed matrix.

Selectivity of the discrimination

The next step was to investigate the units' selectivity for particular calls, in order to assess if the high performance of some units was due to strong discrimination abilities for certain calls of a few particular individuals (i.e., a high selectivity), or to a capacity to discriminate between calls from all individuals (low selectivity). We first calculated the entropy of the distribution of the probabilities of correctly classifying each individual in the confusion matrix (that is, a measure of the randomness of the probability distribution in the diagonal of the confusion matrix): a uniform distribution (that would show an absence of selectivity at the unit level) would give a high entropy value, whereas a heterogeneous distribution (that would be the sign of a higher selectivity for the acoustic characteristics of certain calls) would give a low entropy value. The results are shown in Fig. 5a: interestingly, all the most discriminant units have very low selectivity in their discrimination (i.e. very high entropy values). To further investigate this selectivity at the unit level, we calculated a mean percentage of correct classification (PCC) taking only the best-discriminated individual from each confusion matrix. Our goal was to assess if the best-discriminated individuals were found for the highly discriminant units, or if the high discrimination performance of these units was merely due to a good (but not maximal) discrimination of each individual. As can be seen on Fig. 5b, the highest PCC at the level of the individual were found for the best units overall.

Thus, the units showing the best overall performance in discriminating between individuals are also the ones showing the best discrimination performance per individual, their low selectivity indicating additionally that they perform well for each individual.

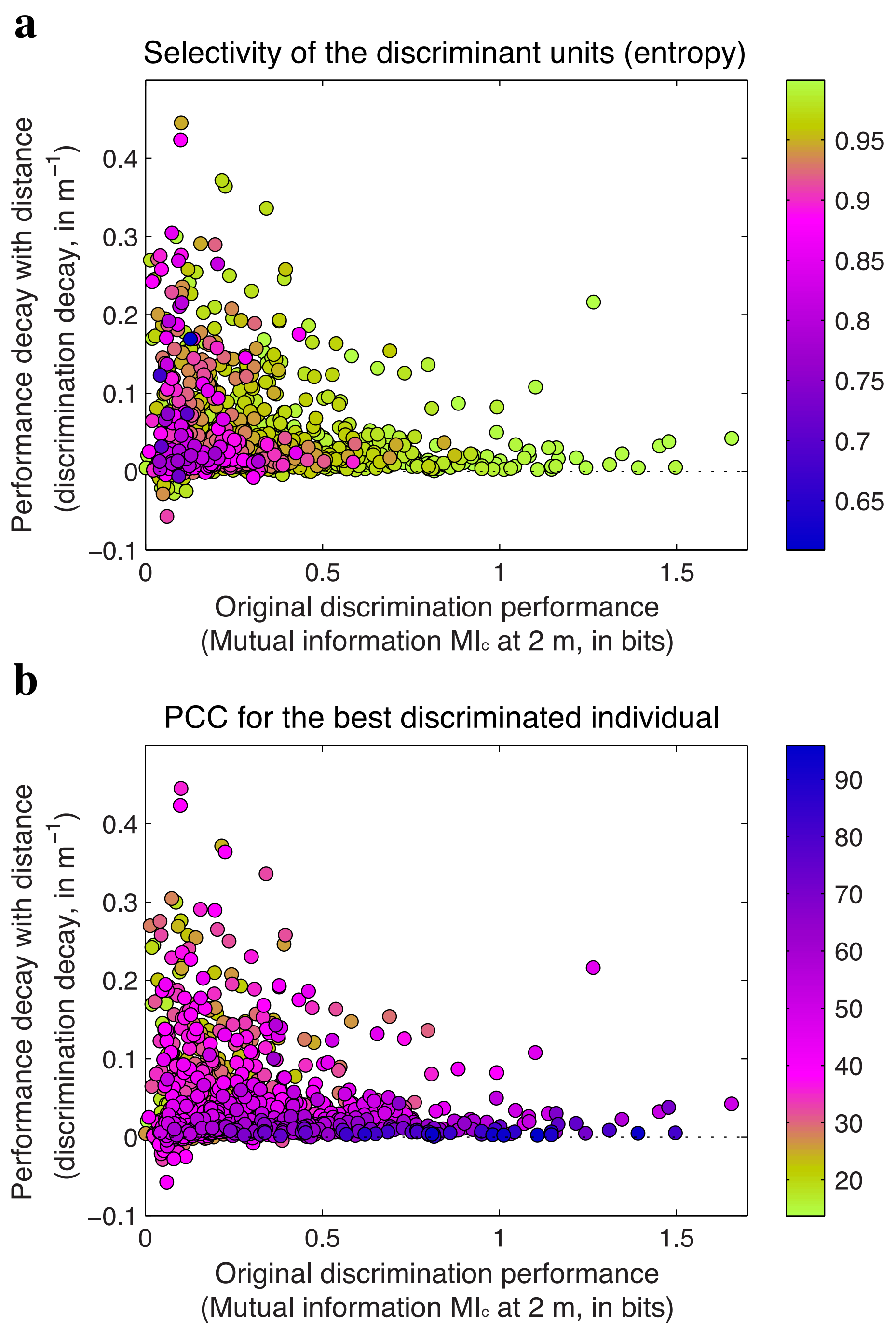


Fig. 5: Between-individuals discrimination performance and selectivity. Both figures show all discriminant units plotted in the same space as in Figure 3. Fig. 5**a** shows on the color scale the mean entropy across distances of the probabilities of correctly classifying each individual in the confusion matrix, that is, a measure of the unit's selectivity for the calls or particular individuals. High entropy values (in green) show low selectivity at the unit level. Fig. 5**b** represents mean percentages of correct classification taking only into account the best-discriminated individual in the confusion matrices across all distances. The probability of correct classification for this individual in the confusion matrix was multiplied by the total number of individuals (4 or 8) so as to obtain a percentage with a maximal value of 100%, and thus compare the values obtained for subjects tested with 4 or 8 different individuals.

Temporal resolution of the neural code

As explained in the methods, the time constant used to calculate the differences between spike patterns (and thus the MI_c) was chosen so as to maximize the discrimination performance of each neuron. Seven values were tested for time constants: 1, 3, 5, 10, 30, 50 and 100 ms. Fig. 6a shows the distribution of the time constant values chosen across all discriminant units, and indicates that the time constants yielding maximal discrimination were small (5 ms and 10 ms constants represented 53.7% of all units). Looking closer at the highly discriminant units only by taking the 139 neurons that had a mean PCC across distances above 20% over chance level (Fig. 6b), we found that the most represented constants were the smaller ones (from 3 to 10 ms), with the 5 ms constant representing 43.9% of all selected units. Indeed, discrimination performance (measured as PCCs) was inversely correlated to the size of the time constants (correlation coefficient $r = -0.25$, $p < 10^{-4}$). Thus, the most discriminant units used a neural code that relied more heavily on spike patterns than the firing rate. This dependence on spike patterns does not of course exclude the effect of the overall spike rate. Indeed, we also found a significant correlation between discrimination and rate measured as a raw firing rate ($r = 0.35$, $p < 10^{-4}$) or as a z-score ($r = 0.68$, $p < 10^{-4}$): higher firing rates and in particular more reliable firing with respect to background yielded higher discrimination performances (Fig. 7). Thus, the highest discrimination was obtained for neurons that have high mean firing during sounds relative to background *and* that modulate this firing rate in time to obtain distinct spike patterns for distinct sounds.

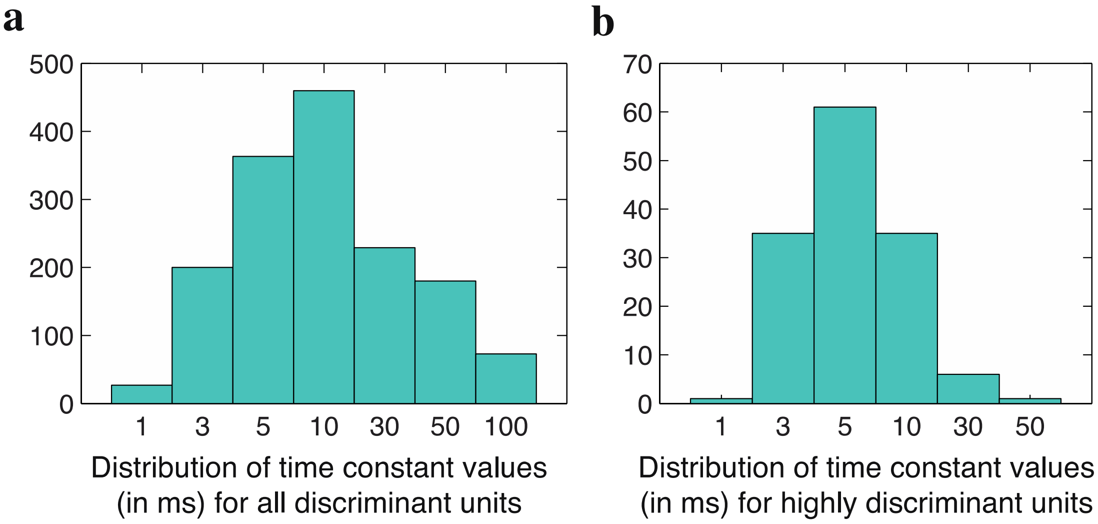


Fig. 6: distribution of the time constants (in ms) used for all discriminant units (a) and for highly discriminant units (b). Highly discriminant units were defined as having a mean PCC above 20% (n = 139). For each unit the time constant yielding the highest discrimination was chosen from a set of 7 values ranging from 1 to 100 ms.

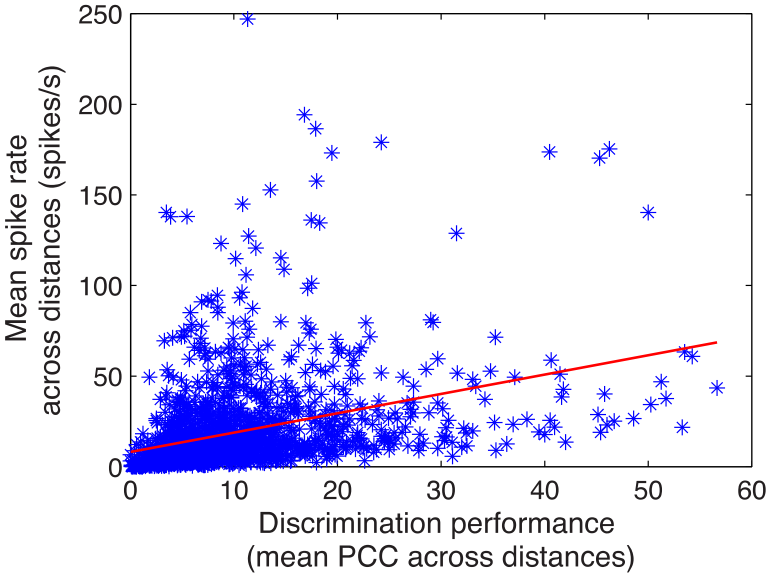


Fig. 7: Mean spike rate across distances (in spikes/s) as a function of the discrimination performance (measured by the mean PCC across distances) for all discriminant units. The regression line calculated using all points is represented in red.

Anatomical localization of the highly-discriminating units

Focusing on the discriminating units showing a mean PCC above chance level across all distances superior to 45%, we found that over 16 units, 11 were located in the field L complex and 2 in the secondary auditory areas (1 in NCM, 1 in CMM); we could not locate the last 3 units. For the units located in the field L complex, we found 2 in L2b, 1 in L3 and 3 in the sub-field L; the 5 field L units remaining were undetermined. This histological analysis needs to be performed more thoroughly to assess if this trend is confirmed when looking at a bigger population of discriminant neurons.

Effect of intensity in the propagation-induced degradation of neural discrimination

In the 'Con+Syn' protocol, each subject was challenged with two sets of stimuli in order to disentangle the overall effect of sound propagation through natural environment from the mere effect of the decrease in sound intensity on the units' responses. The Syn stimuli mimic the intensity decrease of the Con stimuli at the equivalent distance (from 16 to 256 m), but with the same high signal to noise ratio obtained from the call recorded at 2 m. To illustrate this, we show on Fig. 8 the responses of a representative unit to the Con and Syn versions of the same call, for all distances. The decay curve and confusion matrices show only minor differences in this unit's ability to discriminate individuals with these two sets of calls.

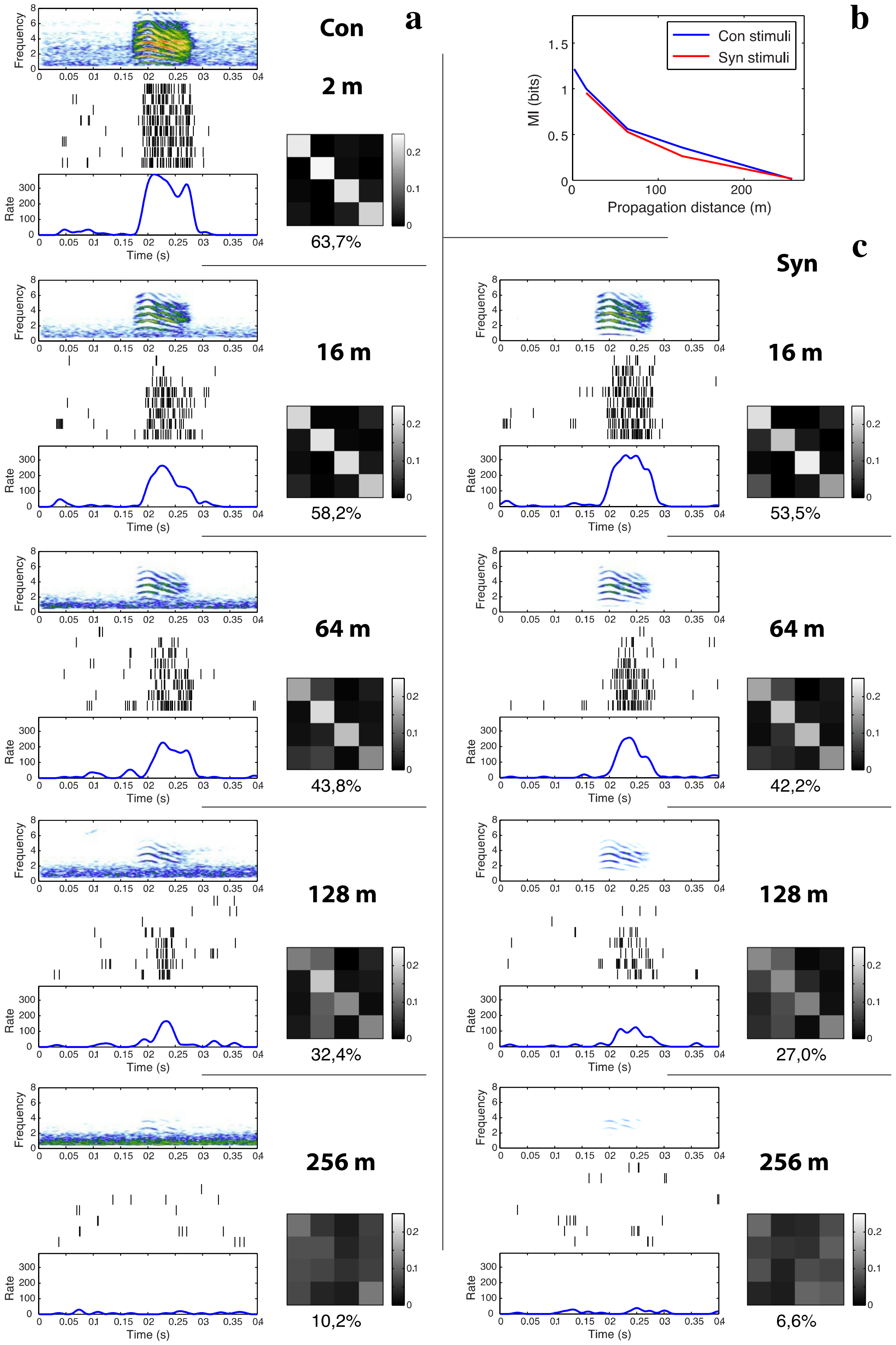


Fig. 8: Comparison of the responses of the same unit to the Con and Syn versions of the same call. The decay curves for each stimulus type are shown for this unit (**b**). For each propagation distance, the neural responses are shown for the Con (**a**) and the Syn stimuli (**c**). As in Fig. 2, for each distance the spectrogram of the stimulus is shown (Frequency in kHz), as well as the spike trains (Rate in spikes/s) for the 8 trials, the PSTH and the confusion matrix with the PCC over the chance level.

This result was also observed at the population level. Fig. 9 shows that the mean MI values for each distance for the Con and Syn stimuli are strikingly equivalent (paired t-tests showed no significant difference, for every distance). In order to further compare the responses to the Con and Syn stimuli, we also calculated the difference between the discrimination decay for the Syn and the Con stimuli (decay for Syn - decay for Con) for each unit; this difference is not significant (t-test: $p = 0.09$). The same test taking only highly discriminant units (for which mean PCC > 20%) was even less significant ($p = 0.88$).

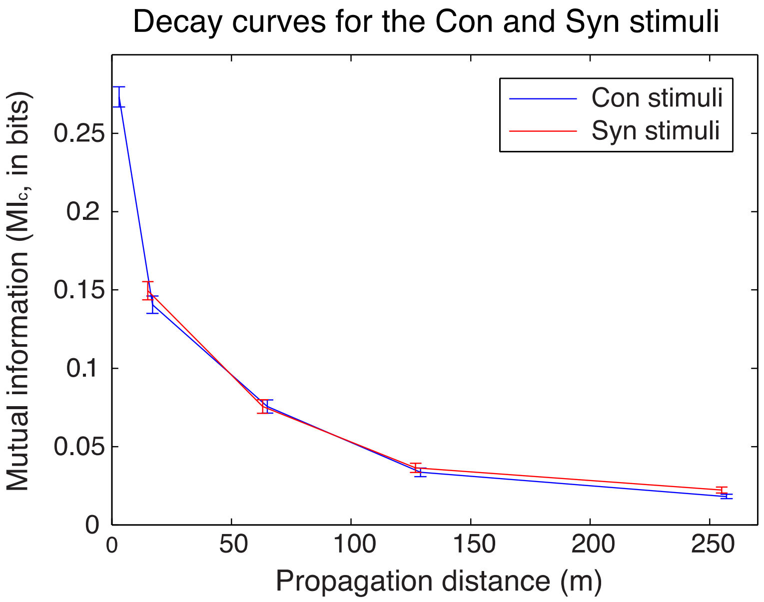


Fig. 9: discrimination performance for Con and Syn stimuli. The plot shows the mean decay curves for the responses to the Con and Syn stimuli at all distances. Error bars represent the standard error of the mean, and jitter on the x-axis has been added for better visibility.

Although the difference in discriminability between Con and Syn stimuli is non-significant overall, we found a small number of neurons for which the performance in discriminability for the Syn stimuli was substantially higher than for the Con stimuli. The example shown on Fig. 10 illustrates that this difference is mainly due to the unit's sensitivity to background noise for the Con stimuli.

Thus, it appears that for most discriminant units, the degradation of the units' performance with distance is mostly explained by the mere decrease in intensity rather than by the increase in signal degradation. We found some units however for which the discrimination performance was higher for the Syn stimuli than the Con, and in that case their higher sensitivity to background noise explained the fact that the Syn stimuli, with their higher signal-to-noise ratio (SNR), yielded better performances.

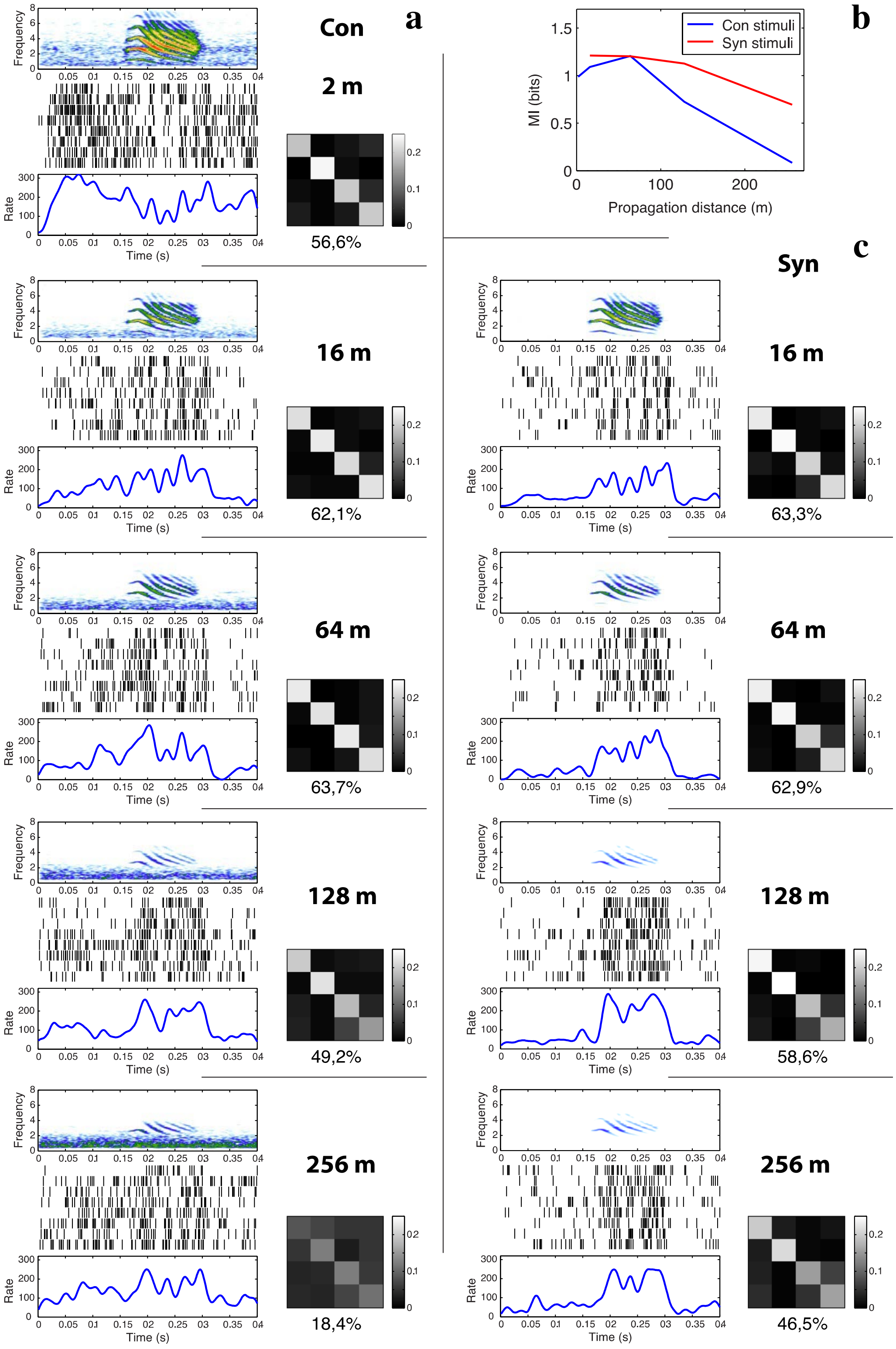


Fig. 10: comparison of the responses of the same unit to the Con and Syn versions of the call, second example. Here we show the responses for both versions of the same call for a unit for which the discrimination performance was notably higher for the Syn than for the Con stimuli. The decay curves for each stimulus type are shown for this unit (**b**). For each propagation distance, the neural responses are shown for the Con (**a**) and the Syn stimuli (**c**). As in Fig. 2, for each distance the spectrogram of the stimulus is shown (Frequency in kHz), as well as the spike trains (Rate in spikes/s) for the 8 trials, the PSTH and the confusion matrix with the PCC over the chance level.

Effect of sex on the discrimination performance

We tested the effect of sex on the discrimination performance, both at the level of the stimulus and at the level of the subject, using the Con stimuli. Of all the discriminant units tested on both male (n = 913) and female subjects (n = 619), 784 were tested only with male stimuli and 748 only with female stimuli. In order to take into account both the sex of the subject and the sex of the individual used as stimulus, we performed a linear mixed-effect model (using the lmer function on R) with these predictors, using the mean PCC over the chance level as the response variable and taking into account the effect of the subject as a random factor. The statistical analysis showed a significant effect of the sex associated to the stimulus on the discrimination performance ($\chi^2(1) = 54, p < 10^{-4}$), but no effect of the subject's sex as a main effect. However, the interaction between both was significant ($\chi^2(1) = 15.5, p < 10^{-4}$). Fig. 11 illustrates the effect of both factors as explained by the model, and shows a surprising trend for the male subjects: while females performed similarly for stimuli of both sexes, males showed a striking difference in their discrimination performance, yielding higher PCCs than females when discriminating between male calls and lower PCCS than females when discriminating between female calls. Fig. 12 shows the discriminant units' performance depending on the sex of the individuals used as stimuli for each unit or the sex of the tested subject, and further illustrates the fact that the best PCCs were obtained when the units discriminated between male stimuli, and that male subjects were more proficient at discriminating between male stimuli.

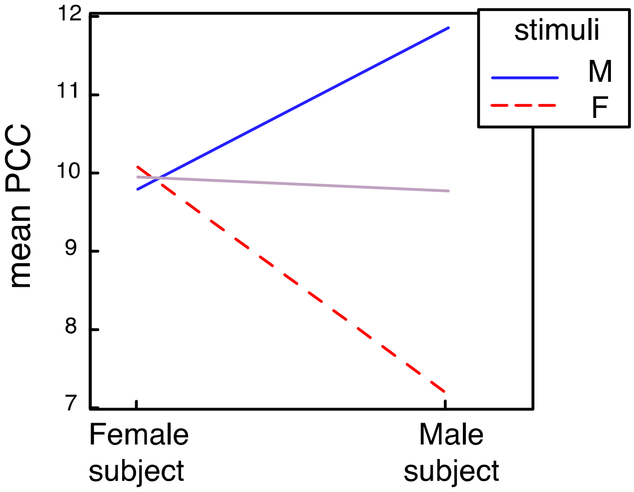


Fig. 11: Differential effects of the subject's sex or the sex associated to the stimuli, as explained by the model. The mean PCC is shown for female subjects discriminating between stimuli from both sexes, and likewise for the male subjects. Males are represented in blue, females in red. The pale line in the middle is the mean PCC across stimulus type for each subject's sex.

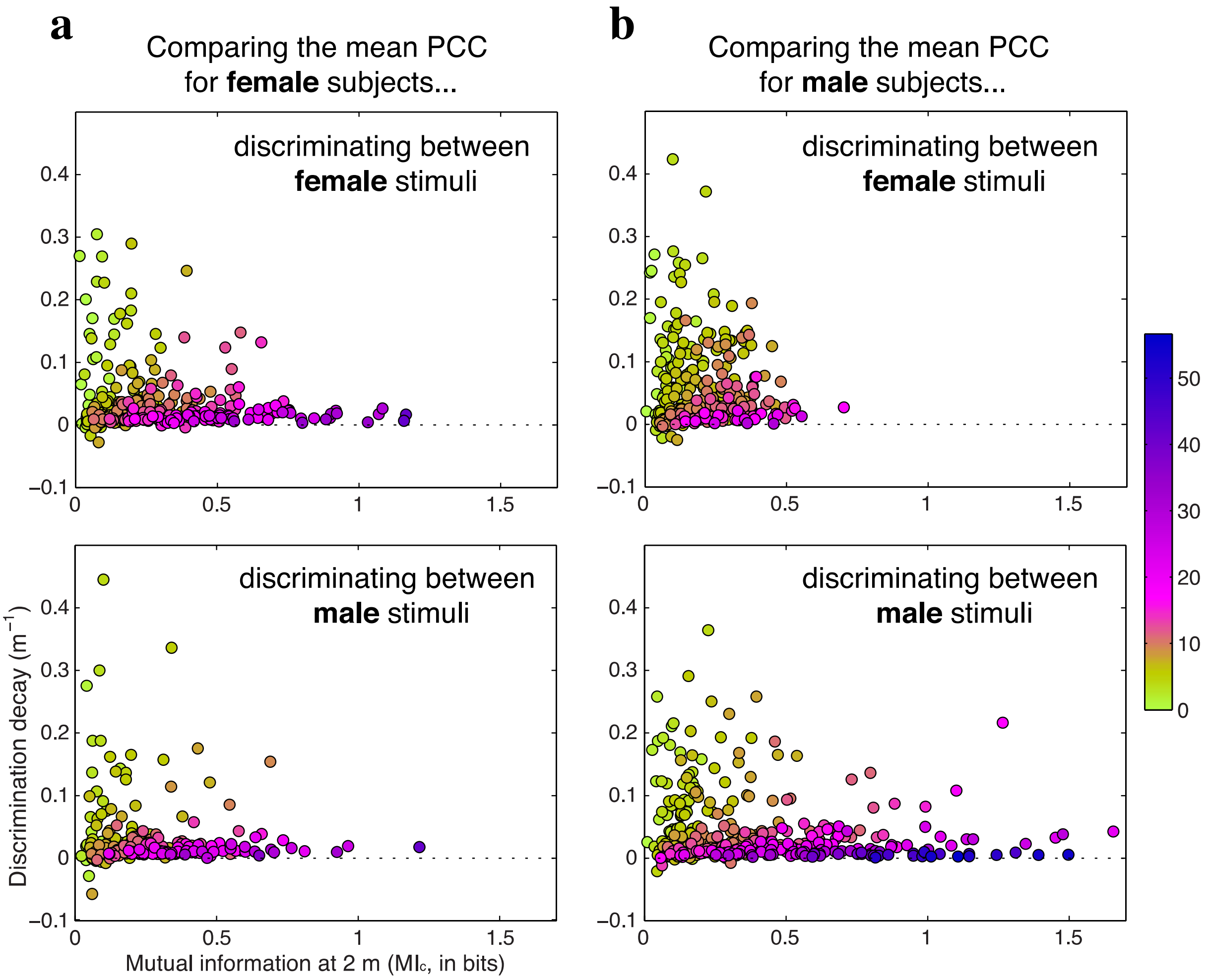


Fig. 12: Discrimination performance for each sex of the subject and the stimuli. (**a**): Both figures represent the discriminant units plotted in the same space as in *Fig. 3*, showing the discrimination performance of female subjects when discriminating between female and between male stimuli. (**b**): the discriminant units are plotted in the same space, showing the discrimination performance of male subjects when discriminating between female and between male stimuli. The discrimination performance is shown on the color scale using the mean PCCs over chance level calculated from the responses to the Con stimuli.

DISCUSSION

We found a population of neurons in the avian auditory forebrain that discriminate individual voices with various degrees of propagation-induced degradation without prior familiarization or training. The discrimination of individual voices at long distance is a behaviorally relevant problem that is encountered in the wild by our model species, the zebra finch, and our use of stimuli recorded in real-life conditions ensured that our study was closely related to this biological task. The best performing neurons were able to discriminate individuality in calls as far as 250 m, which represents a very high degradation level for these vocalizations in regard to the signal-to-noise ratio (SNR). Interestingly, the neurons showing the highest degree of individual discrimination were also the least selective; in other words they showed high discrimination performances towards all individuals that were presented. Furthermore, we found that most of these neurons could discriminate individuality using the same (or very similar) spike pattern code regardless of distance, indicating that they did not only categorize different individuals as different entities at each distance but that they could categorize individual identity regardless of propagation distance.

It should be noted that this discrimination ability was found in birds that had been born and raised in captivity, and had therefore no experience in processing and interpreting vocalizations that had been propagated at distances greater than ~ 5 m. Thus, we deduce that the neural mechanisms needed to extract higher-order information (i.e., detecting the individual identity in a number of calls from different birds) from naturally degraded calls exist without prior experience of propagation-induced call degradation or prior familiarization with the tested calls. In order to discriminate the propagated calls used in our experiments, neurons had to deal with multiple modifications in the sounds. First, the intensity of the signal and the SNR decreased with increasing distances. Second, the spectro-temporal structure of the signal was modified: in the propagated calls used as stimuli, the frequency band with a SNR greater than 0 dB was progressively reduced with distance to a narrower band around 3.5 kHz; the spectral modulations such as those observed in harmonic stacks lost their sharpness and, similarly, the temporal envelope of the sound was smeared in time (Mouterde et al. *submitted*).

Previous studies had also described invariance properties of discriminant neurons, either to intensity (using single exemplars of songs from two or five individuals; Billimoria et al., 2008), or to background noise (Moore et al., 2013; Schneider and Woolley, 2013). These last two studies, using either unfamiliar songs in anesthetized subjects or learned songs in awake animals respectively, pointed out that training and behavioral relevance were not critical for the neural extraction of songs from background noise. Here, using calls that have been degraded in a natural environment and testing multiple exemplars of calls for each individual, we also show that previous training is not critical for individual discrimination in degraded calls for both sexes. Moreover, our results demonstrate that in order to perform this discrimination task in a real-life auditory scene analysis, neurons need to show invariance not only to sound level *per se* but also to SNR as used in the broad sense, i.e. taking into account the intensity of the signal relative to the background noise as well as the degradation of the signal's spectral and temporal structure due to propagation. Indeed, the best performing neurons in our study show a certain degree of invariance towards intensity (their discrimination performance decreasing with distance but staying well over chance even as far as 250 m) as well as a strong invariance towards the progressive decrease in the signals' SNR (as evidenced by the equivalent discrimination performance between the Con and Syn stimuli, that is, between the naturally degraded calls and the modified calls for which the sound level decreased with distance but the SNR was constant).

While critical neural substrate for the analysis of auditory scenes has been found in field L (intensity invariant neurons, as explained above), previous studies have highlighted the importance of the secondary auditory areas (NCM and CM) in processing higher-order features such a behavioral significance, and their role in learned auditory discrimination such as individual recognition (Chew et al., 1996; Gentner, 2004; Pinaud and Terleph, 2008). Moreover, the fact that noise invariant neurons were found in the NCM may have oriented our search for neurons performing auditory scene analysis while discriminating individuality in the secondary auditory area. The fact that, in our study, the majority of the best performing neurons was found in field L demonstrates that this complex task can be performed in the primary auditory area. Although the localization of the most discriminant units is still at a preliminary stage in this study, it is still interesting to compare these results with a previous study about the spectro-temporal coding properties of field L in zebra finches: using a stimulus based on properties of natural sounds, Nagel and Doupe (2008) found that a majority of neurons in field L was narrowly tuned and sensitive only to a subset of possible acoustic features. While we do not present any results about spectro-temporal coding in this manuscript, the fact that we found neurons able to encode individual identity robustly regardless of propagation-induced degradations suggests the existence of more broadly tuned neurons in field L than what was previously thought. We suggest that spectro-temporal processing required for individual discrimination even between degraded calls is present in the primary auditory cortex and might sub-serve an all-purpose and spontaneous voice detection task. Processing in secondary auditory areas could then develop following the bird's experience for the learned recognition of specific individuals (e.g., learning to recognize its mate at a distance). In the associative model of individual vocalizations recognition (Gentner, 2004), this population of neurons we described would take part in a bottom-up process creating a potential for individual discrimination, while top-down processes such as attentional or motivational mechanisms would develop this potential through perceptual plasticity. Another kind of discrimination ability has been shown in rhesus macaques (*Macaca mulatta*), for which the ventrolateral prefrontal cortex is involved in the categorization of socially meaningful but unfamiliar communication calls (Gifford et al., 2005).

Another interesting finding is that male calls were significantly better discriminated than female calls. This corroborates previous studies that showed that the males' individual signature was stronger than the females' (Zann, 1984), and even more so when propagated at long distance

(Mouterde et al. *submitted*). Our results also showed that neurons recorded in males were more proficient at discriminating between male stimuli than female stimuli. To our knowledge, this is the first time that such a finding is reported, and both the magnitude of this effect in males and the lack of such finding in females are very surprising to us. While it would be insightful to further investigate this phenomenon, we attempt to explain it by the fact that male zebra finches have been shown to be more capable than females for processing particular acoustic features in calls (Vicario et al., 2001), and it has been suggested that because the males learn their distance call (and their call being very close to a syllable type of their song; Zann, 1996), there might be a functional or even a causal relationship between the males' ability to produce learned features and their tendency to categorize these features, and thus better discriminate between them (Vicario et al., 2001). This, combined with the stronger individual signature of males, may be a beginning of explanation as to why males are so much more proficient at discriminating between male calls, but at this stage we still fail to explain why males showed worse discrimination for female calls than did females and why females did not show enhanced discrimination for male calls relative to female calls.

Ultimately, one of the main goals in brain research is to link the processes described at the neural level with corresponding behavioral manifestations. This attempt to integrate neural discrimination and behavior is difficult since we cannot assume that animals use the information present in particular neurons without specifically testing perception (Nagel et al., 2011). An interesting step in our understanding of these processes is that behavioral performance has been found to be more strongly correlated with the performance of the best neurons rather than the average neuron (Wang et al., 2007). Using behavioral tests, we have found in a previous study that female zebra finches were able to discriminate between the calls of male zebra finches without training at up to 128 m, and up to 256 m when they were able to learn from one day to the next (Mouterde et al. *in revision*). These results were obtained with a forced-choice conditioning task using fasted birds with food as reward, which would lead the subjects to express their best discrimination performance, stripped from any social effect (e.g., compared to a preference test). Thus, one could argue that in this case, the behavioral tests showed lower discrimination performance than what was found for the best neurons in this study, which presented an ability for spontaneous individual discrimination as far as 256 m. However, only females were tested in the behavioral experiments, and following our findings that males showed higher discrimination performances, it is possible that when tested in this operant setup they would show such abilities at further distances without prior training. Additionally, training subjects to recognize degraded individual vocalizations in this type of task before recording their neural responses to the same signals would give valuable insight into the mechanisms of individual voice recognition in an auditory scene as well as the underlying learning processes and memory formation. Indeed, there are still too few experiments where comparisons are performed between neurometric and psychometric functions, and so far, only a small number of these experiments have been performed within the same subjects (Dong et al., 2011; Huetz et al., 2011). It would be interesting to compare the results of such experiments to the ones presented here using naive subjects.

Our study tested neural discrimination in anesthetized birds, and one could object that these recording conditions are further from the reality of a bird discriminating voices in an auditory scene in the wild than having awake animals during recordings. This is a fair point, and it would indeed be interesting to test this question with awake animals, but we emphasize the fact that such a study testing spontaneous discrimination in neurons would necessitate less top-down effects, such as selective attention, that would be depressed with anesthesia. Furthermore, comparative studies on the effect of anesthesia in the midbrain (Schumacher et al., 2011) and the primary forebrain (Narayan, 2006) of songbirds found no differences in discriminability performances and spectral tuning of single neurons among urethane-anesthetized and unanesthetized responses to communication vocalizations in zebra finches. While urethane depressed intrinsic neural excitability in the midbrain, no significant difference in firing rates could be found for field L neurons (Narayan, 2006). Thus, and bearing in mind that the effect of anesthesia can vary greatly between species (Gaucher et al., 2013), we can safely assume that the neural responses we recorded in our model species were similar to what we would have found using awake animals.

Concerning the neural coding of individual degraded voices, our finding that taking into account the temporal patterns of discharges yielded high discrimination performances while the average firing rate did not reliably predict these performances corroborates numerous studies claiming that neurons do not show a selectivity for the vocalizations in terms of firing rate but, instead, use a spike timing strategy to encode natural vocalizations, especially in field L (Gaucher et al., 2013; Huetz et al., 2011; Schumacher et al., 2011; Wang et al., 2007). However, as Theunissen and Miller (1995) pointed out, when spike timing is related to the timing of specific features in the stimulus, the neural code should be labeled as a rate code with a fast temporal scale. Here we found that the optimal timescales for neural discrimination of degraded voices were around 5 ms, which is consistent, if somewhat lower, with what was found previously for neural discrimination of conspecific song in the field L of zebra finches (Narayan, 2006). An interesting remark about the coding properties of neurons solving difficult signal processing problems is that these neurons are usually found to have sparse firing patterns; this sparse coding has been found in particular for neurons performing auditory scene analysis (Asari, 2006; Billimoria et al., 2008; Schneider and Woolley, 2013), where they have been showed to be more invariant to noise or to intensity. Indeed, according to Gaucher *et al.* (2013), temporal sparse code of auditory cortex neurons can be considered as a first step for generating high-level representations of communication sounds independent of the acoustic characteristic of these sounds. In the present study, on the one hand, both the lack of individual selectivity for highly discriminant neurons and the positive correlation between discrimination performance and firing rate does not appear to support a sparse coding hypothesis. On the other hand, it is possible that these highly discriminant units are particularly selective for calls relative to other sounds and could be sparse in their representation in this sense. The next step in understanding the neural coding of vocalizations is thus both to measure responses to other stimuli and to assess to which extent the temporal spike patterns are linked to the acoustic structure of the stimuli, and which part could be attributed to higher-order representations such as behavioral meaning (Huetz et al., 2011). In other words, does the discrimination ability that we found in auditory neurons depend exclusively and in a graded fashion on the acoustical differences between calls, or is there a higher level of integration yielded by a non-linear and categorical group of voices? Investigating this question with the use of spectro-temporal receptive fields (STRFs) or other encoding models would complete the decoding approach used here (Theunissen et al., 2000). These approaches would help understand how the neural coding initially based on the vocalization's acoustic properties becomes more integrated and potentially categorical as one ascends the auditory pathway.

Other research approaches offer exciting challenges in the study of how the brain processes auditory scenes: exploring neural networks rather than single neurons (Quian Quiroga and Panzeri, 2009) or the plasticity of perception and the impact of top-down processes in neural coding (Chapuis and Wilson, 2011; Dahmen and King, 2007; Gentner, 2004; Menardy et al., 2012) will help us better understand how the complex task of scene analysis is performed by the brain.

REFERENCES

Amin, N., Gastpar, M. and Theunissen, F. E. (2013). Selective and Efficient Neural Coding of Communication Signals Depends on Early Acoustic and Social Environment. *PLoS ONE* **8**, e61417.

Appeltants, D., Gentner, T. Q., Hulse, S. H., Balthazart, J. and Ball, G. F. (2005). The effect of auditory distractors on song discrimination in male canaries (*Serinus canaria*). *Behav. Processes* **69**, 331–341.

Asari, H. (2006). Sparse Representations for the Cocktail Party Problem. *J. Neurosci.* **26**, 7477–7490.

Aubin, T. and Jouventin, P. (2002). How to vocally identify kin in a crowd: The penguin model. In *Advances in the Study of Behavior*, pp. 243–277. Elsevier.

Bailey, D. J., Rosebush, J. C. and Wade, J. (2002). The hippocampus and caudomedial neostriatum show selective responsiveness to conspecific song in the female zebra finch. *J. Neurobiol.* **52**, 43–51.

Billimoria, C. P., Kraus, B. J., Narayan, R., Maddox, R. K. and Sen, K. (2008). Invariance and Sensitivity to Intensity in Neural Discrimination of Natural Sounds. *J. Neurosci.* **28**, 6304–6308.

Boumans, T., Vignal, C., Smolders, A., Sijbers, J., Verhoye, M., Van Audekerke, J., Mathevon, N. and Van der Linden, A. (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.

Bregman, A. S. (1993). Auditory scene analysis: hearing in complex environments. In *Thinking in sound: the cognitive psychology of human audition* (ed. McAdams, S. E. and Bigand, E. E.), Oxford.

Butterfield, P. A. (1970). The pair bond in the zebra finch. In *Social behaviour in birds and mammals: Essays on the social ethology of animals and man*, pp. 249–278. London: J. H. Crook.

Chapuis, J. and Wilson, D. A. (2011). Bidirectional plasticity of cortical pattern recognition and behavioral sensory acuity. *Nat. Neurosci.* **15**, 155–161.

Charrier, I., Mathevon, N. and Jouventin, P. (2003). Vocal signature recognition of mothers by fur seal pups. *Anim. Behav.* **65**, 543–550.

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

Chicharro, D., Kreuz, T. and Andrzejak, R. G. (2011). What can spike train distances tell us about the neural code? *J. Neurosci. Methods* **199**, 146–165.

Dahmen, J. C. and King, A. J. (2007). Learning to hear: plasticity of auditory cortical processing. *Curr. Opin. Neurobiol.* **17**, 456–464.

Dong, C., Qin, L., Liu, Y., Zhang, X. and Sato, Y. (2011). Neural Responses in the Primary Auditory Cortex of Freely Behaving Cats While Discriminating Fast and Slow Click-Trains. *PLoS ONE* **6**, e25895.

Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans.* Society for Industrial and Applied Mathematics.

Forrest, T. G. (1994). From sender to receiver: propagation and environmental effects on acoustic signals. *Integr. Comp. Biol.* **34**, 644–654.

Fortune, E. S. and Margoliash, D. (1992). Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (*Taenopygia guttata*). *J. Comp. Neurol.* **325**, 388–404.

Gastpar, M. C., Gill, P. R., Huth, A. G. and Theunissen, F. E. (2010). Anthropic Correction of Information Estimates and Its Application to Neural Coding. *IEEE Trans. Inf. Theory* **56**, 890–900.

Gaucher, Q., Huetz, C., Gourévitch, B., Laudanski, J., Occelli, F. and Edeline, J.-M. (2013). How do auditory cortex neurons represent communication sounds? *Hear. Res.* **305**, 102–112.

Gentner, T. Q. (2004). Neural Systems for Individual Song Recognition in Adult Birds. *Ann. N. Y. Acad. Sci.* **1016**, 282–302.

Gentner, T. Q. and Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature* **424**, 669–674.

Gifford, G. W., MacLean, K. A., Hauser, M. D. and Cohen, Y. E. (2005). The Neurophysiology of Functionally Meaningful Categories: Macaque Ventrolateral Prefrontal Cortex Plays a Critical Role in Spontaneous Categorization of Species-Specific Vocalizations. *J. Cogn. Neurosci.* **17**, 1471–1482.

Grace, J. A., Amin, N., Singh, N. C. and Theunissen, F. E. (2003). Selectivity for Conspecific Song in the zebra finch auditory forebrain. *J. Neurophysiol.* **89**, 472–487.

Hauber, M. E., Cassey, P., Woolley, S. M. N. and Theunissen, F. E. (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.

Hill, D. E. (1979). Orientation by jumping spiders of the genus Phidippus (Araneae: Salticidae) during the pursuit of prey. *Behav. Ecol. Sociobiol.* **5**, 301–322.

Hsu, A., Borst, A. and Theunissen, F. E. (2004a). Quantifying variability in neural responses and its application for the validation of model predictions. *Netw. Comput. Neural Syst.* **15**, 91–109.

Hsu, A., Woolley, S. M. N., Fremouw, T. E. and Theunissen, F. E. (2004b). Modulation Power and Phase Spectrum of Natural Sounds Enhance Neural Encoding Performed by Single Auditory Neurons. *J. Neurosci.* **24**, 9201–9211.

Huetz, C., Gourévitch, B. and Edeline, J.-M. (2011). Neural codes in the thalamocortical auditory system: From artificial stimuli to communication sounds. *Hear. Res.* **271**, 147–158.

Latinus, M., McAleer, P., Bestelmeyer, P. E. G. and Belin, P. (2013). Norm-Based Coding of Voice Identity in Human Auditory Cortex. *Curr. Biol.* **23**, 1075–1080.

Lewicki, M. S., Olshausen, B. A., Surlykke, A. and Moss, C. F. (2014). Scene analysis in the natural environment. *Front. Psychol.* **5**,.

Logothetis, N. K. and Sheinberg, D. L. (1996). Visual Object Recognition. *Annu. Rev. Neurosci.* **19**, 577–621.

Mathevon, N., Aubin, T., Vielliard, J., da Silva, M.-L., Sèbe, F. and Boscolo, D. (2008). Singing in the rain forest: how a tropical bird song transfers information. *PLoS ONE* **3**, e1580.

Menardy, F., Touiki, K., Dutrieux, G., Bozon, B., Vignal, C., Mathevon, N. and Del Negro, C. (2012). Social experience affects neuronal responses to male calls in adult female zebra finches. *Eur. J. Neurosci.* **35**, 1322–1336.

Moore, R. C., Lee, T. and Theunissen, F. E. (2013). Noise-invariant Neurons in the Avian Auditory Cortex: Hearing the Song in Noise. *PLoS Comput. Biol.* **9**, e1002942.

Nagel, K. I. and Doupe, A. J. (2008). Organizing Principles of Spectro-Temporal Encoding in the Avian Primary Auditory Area Field L. *Neuron* **58**, 938–955.

Nagel, K., Kim, G., McLendon, H. and Doupe, A. (2011). A bird brain's view of auditory processing and perception. *Hear. Res.* **273**, 123–133.

Narayan, R. (2006). Distinct Time Scales in Cortical Discrimination of Natural Sounds in Songbirds. *J. Neurophysiol.* **96**, 252–258.

Narayan, R., Best, V., Ozmeral, E., McClaine, E., Dent, M., Shinn-Cunningham, B. and Sen, K. (2007). Cortical interference effects in the cocktail party problem. *Nat. Neurosci.* **10**, 1601–1607.

Okanoya, K. and Dooling, R. J. (1987). Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. *J. Comp. Psychol.* **101**, 7–15.

Pinaud, R. and Terleph, T. A. (2008). A songbird forebrain area potentially involved in auditory discrimination and memory formation. *J. Biosci.* **33**, 145–155.

Quian Quiroga, R. and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* **10**, 173–185.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107.

Ribeiro, S., Cecchi, G. A., Magnasco, M. O. and Mello, C. V. (1998). Toward a Song Code. *Neuron* **21**, 359–371.

Rossum, M. C. W. van (2001). A novel spike distance. *Neural Comput.* **13**, 751–763.

Schneider, D. M. and Woolley, S. M. N. (2013). Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes. *Neuron* **79**, 141–152.

Schnitzler, H.-U. and Flieger, E. (1983). Detection of oscillating target movements by echolocation in the Greater Horseshoe bat. *J. Comp. Physiol. A* **153**, 385–391.

Schumacher, J. W., Schneider, D. M. and Woolley, S. M. N. (2011). Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. *J. Neurophysiol.* **106**, 500–514.

Sen, K., Theunissen, F. E. and Doupe, A. J. (2001). Feature analysis of natural sounds in the songbird auditory forebrain. *J. Neurophysiol.* **86**, 1445–1458.

Stopfer, M., Jayaraman, V. and Laurent, G. (2003). Intensity versus identity coding in an olfactory system. *Neuron* **39**, 991–1004.

Theunissen, F. and Miller, J. P. (1995). Temporal encoding in nervous systems: A rigorous definition. *J. Comput. Neurosci.* **2**, 149–162.

Theunissen, F. E. and Shaevitz, S. S. (2006). Auditory processing of vocal sounds in birds. *Curr. Opin. Neurobiol.* **16**, 400–407.

Theunissen, F. E., Sen, K. and Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci.* **20**, 2315–2331.

Vates, E. G., Broome, B. M., Mello, C. V. and Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (Taenopygia guttata). *J. Comp. Neurol.* **366**, 613–642.

Vicario, D. S., Naqvi, N. H. and Raksin, J. N. (2001). Sex differences in discrimination of vocal communication signals in a songbird. *Anim. Behav.* **61**, 805–817.

Vignal, C., Mathevon, N. and Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature* **430**, 448–451.

Vignal, C., Mathevon, N. and Mottin, S. (2008). Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behav. Processes* **77**, 191–198.

Von der Emde, G. (2004). Distance and shape: perception of the 3-dimensional world by weakly electric fish. *J. Physiol.-Paris* **98**, 67–80.

Wang, L., Narayan, R., Grana, G., Shamir, M. and Sen, K. (2007). Cortical Discrimination of Complex Natural Stimuli: Can Single Neurons Match Behavior? *J. Neurosci.* **27**, 582–589.

Wiley, R. H. and Richards, D. G. (1982). Adaptations for acoustic communication in birds: sound transmission and signal detection. In *Acoustic Communication in Birds* (ed. Kroodsma, D. E. and Miller, E. H.), pp. 131–181. New York.

Woolley, S. M. N., Fremouw, T. E., Hsu, A. and Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat. Neurosci.* **8**, 1371–1379.

Zann, R. (1984). Structural variation in the zebra finch distance call. *Z. Für Tierpsychol.* **66**, 328–345.

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

CONCLUSION AND PERSPECTIVES

IV. CONCLUSION AND PERSPECTIVES

A. The individual signature of zebra finches is very resistant to degradation

The acoustic analysis of the distance calls from male and female zebra finches showed that information about individual identity is still present in calls propagated as far as 256 m in natural conditions. The use of information contained in the whole spectrogram yielded better results than the traditional approach separating spectral and temporal cues, highlighting the importance of the information about frequency modulation, especially at long distance. This greatly explains the difference found between sexes, as the typical structure of male calls includes a downsweep component that is still discernable at 256 m. Previous work had concluded that male calls are more individualized than female calls (Zann, 1984), and the analysis presented in **manuscript 1** confirmed that this advantage at near distances is even more important at further distances.

Since the distance calls of zebra finches are used to communicate at long distance after losing visual contact (Zann, 1996), it is reasonable to hypothesize that these vocalizations may have been subjected to a selective force leading to their congruence with the sound transmission characteristics in their habitat. For best resistance to degradation in an open environment, and following the acoustic adaptation hypothesis, an ideal call should be short, have a large bandwidth and a high frequency modulation rate (Ey and Fischer, 2009). Male zebra finches' distance calls fit this description, but female calls are typically longer and have low frequency modulations. While individual information may still be found in female calls propagated at 256 m, albeit to a lower level than for the male calls, the fact that their structure does not maximize their transmission distance leads to consider other potential effects that would result in this dimorphic acoustic structure. Indeed, in their development, zebra finches from both sexes first develop an innate distance call, but the males "overwrite" their inherited call with a call that is learned from a tutor (Forstmeier et al., 2009). Thus, the sex-specific characteristics of the distance call may be a by-product from song learning, which points out that the main selection factor for the shaping of these vocalizations might not have been the habitat characteristics (Ryan and Kime, 2003). This sex difference in song and call learning could also help explain the fact that neurons in male subjects discriminated better between male calls in the electrophysiological recordings (**manuscript 3**); however, the fact that they discriminated the female calls worse than neurons in female subjects yields unanswered questions, and it would be interesting to investigate this question in further studies.

Such as it is, this difference in individual signatures between sexes may have an impact on the strategies used by the birds to reunite after having been separated visually, and it is possible that mates have adapted their behavior around the higher transmission potential of the male call. Indeed, it was found experimentally that when members of a pair were visually, but not acoustically isolated from one another, males called more than females and were more active (Butterfield, 1970); Zann (1996) accordingly observed that free-living males call more than females. In his book, Zann also cited the unpublished work of A. W. McIntosh, who states that "members of a pair relocate in the following way: the male calls more than the female when separated and moves around actively searching for her; when he comes within earshot she recognizes his call and responds with her own distance call but remains stationary: the male homes in on the female that alters her call rate in response to his distance calls". Thus, members of a pair would possibly rely more on the male distance call to establish first contact after being separated, and such a behavior might result from the stronger individual signature that we found in the degraded calls of males.

The distance calls that I used for the acoustic analysis as well as the behavioral and neurophysiological experiments were propagated in a natural environment, and thus combined most aspects of an open field auditory scene: intensity decrease, degradation of the calls' spectro-temporal structure and ambient noise. I am aware however that the analysis of auditory scenes also encompasses notions of directionality and auditory source separation (Lewicki et al., 2014) that were not investigated in this study. The use of a loudspeaker to broadcast sounds effectively annihilates any spatial information that animals in a natural environment can have access to. In nature, different sounds coming from spatially separated sound sources can help enhance discrimination (Bee, 2008; Dent et al., 2009; Maddox et al., 2012). Thus, while this absence of spatial cues may be a caveat when studying the question of auditory scene analysis as a whole, this probably lead us to underestimate the natural discriminability rather than the opposite. However, while the sound transmitted experiments presented in **manuscript 1** showed that the effects of propagation on the recording site in France are similar to those in the Australian desert with little vegetation, in the same low-wind conditions (thus validating the pertinence of our acoustic analysis), other factors may have a strong impact on sound discrimination. Indeed, transmission conditions may vary greatly, and meteorological events such rain or strong wind can greatly affect signal transmission (Brenowitz, 1982), although the former is not frequent in the subarid regions of Australia where zebra finches live. Furthermore, higher levels of ambient noise, due to the calling of insects for example, have an influence on the birds' optimal frequencies for communication (Klump, 1996). More importantly, a chorus from other conspecific birds would have an even higher masking effect, as the frequency range of such masking noise would overlap the frequency range of the tested vocalizations (Forrest, 1994; Lohr et al., 2003). Here I chose to use degraded calls that would represent the "baseline conditions" of an auditory scene, and I am conscious that the distances that we advertise in the manuscripts as enabling voice discrimination would be reduced in more difficult transmission conditions. Besides, it should be noted that the distance calls in the database were broadcasted in "baseline conditions" too, i.e. at a level of 70 dB SPL at 1 m, which is the typical sound level for this species (Vignal et al., 2008). Had I used stimuli recorded in more constraining conditions, I would also have had to take into account the behavioral adaptations that zebra finches have been shown to demonstrate in such conditions; Cynx *et al.* (1998) showed that zebra finches could indeed increase the amplitude of their calls, up to more than 90 dB SPL, in response to white noise. Assessing to which extent the transmission constraints would influence this kind of behavioral adaptation would prove too difficult to analyze in a systematic way, which is why this choice of "baseline conditions" appears to me as the best option for this study of auditory scene analysis.

B. Zebra finches show impressive capacities for discriminating degraded voices

The discrimination task experiment presented in **manuscript 2** showed that without training, female zebra finches could discriminate between the calls of two unfamiliar males propagated as far as 128 m. In the preference test, the female subjects showed a significant preference for their mate as far as 64 m, but were not tested for their preference at 128 m. While in this latter case the subjects were familiar with the stimuli used (their mate or a familiar male), they had never heard vocalizations propagated at such a distance before, as in the former case, and thus had no prior experience with this level of sound degradation. Following the result obtained without training in the discrimination task, we could infer that these females would have shown a significant preference for their mate at 128 m too, had they been tested with this distance. However, this hypothesis is arguable, as the motivation to operate the task in both setups were very different; while discrimination tests at the behavioral level are intrinsically prone to intra-individual variability, it seems that the motivation for a fasted subject to get access to food is greater than the motivation for an isolated subject to hear its mate voice. It is also possible that females in the preference task would have been able to discriminate between calls propagated at 128 m but would not necessarily have shown a clear preference for their mate, the signal's degradation possibly affecting the attractiveness of the stimuli and consequently the consistency of their responses towards it. Thus, the discrimination task served as a mean to avoid the added behavioral flexibility that could be due to context effects of prior social experience in the preference test, and use a highly motivational setup to test the subjects' limits for discrimination between degraded vocal signatures, at the behavioral level.

At the neural level however (**manuscript 3**), I found that a number of single neurons had the ability to discriminate between individual voices as far as 256 m. The performance level of the most discriminant neurons at this distance was comparable to the maximal performance obtained in the acoustic analysis (**manuscript 1**): for males, the mean percentage of correct classification over the chance level (PCC, a measure of the discrimination performance) obtained from the analysis of calls propagated at 256 m was 42.1% while the PCC at this distance for the most discriminant neuron tested with male calls was 44.1%. Although these PCC correspond to two different assessments of discriminability (statistical discrimination from the calls' acoustical structure *versus* between-calls discrimination ability by neurons), the fact that both values are comparable implies that some neurons are able to take the best advantage of the information still remaining after long-distance propagation. The same comparison for females yielded a mean PCC of 28.3% for the acoustic analysis and a value of 22.3% for the most discriminant neuron tested with female calls. These observations highlight the fact that neurons may be able to retrieve a similar amount of information at long distance than what was theoretically calculated using acoustical features optimized for that distance.

Is it possible that zebra finches that have never had the opportunity to hear propagated calls use a "template" for discriminating degraded calls which is optimized for short, medium and long distances? The comparison of the mean PCCs across all distances in both cases suggests that overall, the acoustic analysis using parameters calculated separately for each distance might give better results than the most discriminant neurons (for males: PCC for acoustic analysis = 61.3%, PCC for most discriminant neuron = 56.6%; for females: 54.2% and 41.8% respectively). Thus, the most discriminant neurons that I found seemed to show lower performances than the optimal acoustic analysis overall, but managed to keep their performance at 256 m at a level that was comparable to the acoustic analysis. This high discrimination at long distance is especially interesting when we consider the nature of the prior information available for computing individual discrimination in both cases. For the algorithm using acoustical features optimized for each distance, the fitting dataset (which accounts for the prior information that the tested calls, or validating dataset, were compared to) contained information about the structure of individual distance calls for each individual, and about the distance considered. For the neurons of my recorded subjects, it is probable that the birds had stored some information about acoustical features that are important for discriminating between the voices of different individuals, through their social experience with other birds in the aviary; however, the fact that the vocalizations used as stimuli were recorded from unknown birds implies that their "templates" would never be as efficient as the ones created by the algorithm using these vocalizations. Furthermore, these birds, born and raised in a single room aviary, never had the experience of hearing propagated calls, as stated earlier. This situation seems closer to the "all-distances" template calculated in the acoustic analysis, which maximized average discrimination: in this case, the algorithm contained information about the structure of individual distance calls for each individual and each distance, but no information about the distance per se. This template gave good results at all distances except 256 m, which corresponds to the type of decay curve that I found for a number of discriminant neurons. For the most discriminant neurons however, their high performance at 256 m seems to imply some prior knowledge about distance. Could the perception of distance-dependent sound degradations have an innate basis? While this question could lead to many interesting conjectures, one should not forget that the comparisons between the acoustic analyses and the actual physiology of neurons should be made with extreme caution, and that the reality of the processes at play is very complex; for example, the fact that the auditory system of zebra finches acts like narrowband filters may have an important effect in decreasing the noise within this system (Forrest, 1994).

Another interesting question raised by this discrimination ability without prior familiarization or training that has been shown at the behavioral and neural levels is the effect of the sex of the subject performing the discrimination. In the electrophysiological experiments (**manuscript 3**) we have shown that males are more proficient at discriminating between male calls. In the preference test (**manuscript 2**) only females were tested, following reports from previous work suggesting that the use of male subjects might not yield reliable results in terms of mate preference, making it difficult to take any assumptions on their abilities to recognize degraded calls based on this preference. This was first stated by Zann (1996), citing two unpublished studies showing that females showed a stronger preference towards their mate's distance call than did the males; these observations were later mitigated by Vignal *et al.* (2004), who showed that males could express a preference for their mate but that their response was highly dependent on audience effects. Only females were tested in the discrimination task too, but for practical reasons this time, as only females had been tested with the apparatus at the time of my experiments. It would indeed be interesting to test males in the same task, with female and male calls, in order to assess if the sex-based differences that were found at the neural level are mirrored at the behavioral level.

C. The role of perceptual learning in increasing the discrimination performance

In the discrimination task experiment presented in **manuscript 2**, female zebra finches showed that they were proficient at discriminating between two previously unknown males propagated as far as 256 m when they had the opportunity to practice with these stimuli at increasing distances. This augmented performance compared to the no-training condition, in subjects that had already been shaped in the discrimination task, suggests that the females enhanced their discrimination abilities through perceptual learning (Gold and Watanabe, 2010). This manifestation of experience-dependent plasticity in the sensory systems, which occurs at the developmental stage as well as throughout adult life (Seitz and Dinse, 2007), has been studied at the behavioral level, showing for example that unreinforced learning of song discrimination occurs rapidly in zebra finches, after 3 h of passive song exposure (Stripling et al., 2003). However, the underlying neural mechanisms remain unclear (Dahmen and King, 2007; Yotsumoto et al., 2008). One of the main issues in this field has been 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 (Seitz and Dinse, 2007; Watanabe et al., 2001). Overall, it is probable that 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 (Dahmen and King, 2007; Polley et al., 2006). In any case, it is still a challenge for scientists to relate the neural and behavioral changes at play during perceptual learning (Gold and Watanabe, 2010), and to identify the changes in neuronal response properties that are causally related to perceptual improvements (Dahmen and King, 2007).

While with the behavioral experiments (**manuscript 2**) we have shown that zebra finch females could not discriminate between the degraded voices of males at 256 m without training, we have found at the neural level (**manuscript 3**) that single neurons in naïve birds had the capacity to perform this task. Keeping in mind that both experiments were not conducted on the same subjects, 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 neural substrate for individual discrimination of degraded vocalizations may therefore be a neural basis for vocal recognition in auditory scenes, which could then develop in the secondary auditory areas following the birds' experience for the learned recognition of specific individuals. In other words, this population of 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. This idea converges with 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. This could explain the fact that the discrimination performance that I found at the neural level was matched at the behavioral level only after training.

The fact that such an impressive discrimination ability was found in single neurons provides interesting insight with regard to the existing literature. Previous work dealing with individual discrimination or recognition and auditory scene analysis have dealt with specific aspects of this question, such as invariance to intensity (Billimoria et al., 2008) and to background noise (Moore et al., 2013; Schneider and Woolley, 2013), or individual recognition of undegraded signals (Chew et al., 1996; Gentner, 2004). Thus, it has been hypothesized that the neural basis for the sensitivity to higher-order features in the sounds could arise from an ensemble of responses within the network, each neuron representing a particular stimulus feature (Quian Quiroga and Panzeri, 2009). The study presented in **manuscript 3** highlights however that single neurons are able to discriminate the degraded individual signature of conspecific calls, using the same code (or almost the same code) for each individual regardless of distance. Such discrimination performance requires a high degree of invariance to decreasing signal-to-noise ratio and sound intensity, as well as the ability to extract parameters encoding individual identity in calls for which we have seen in **manuscript 1** that the most individualized features are not the same depending on the propagation distance. Thus, while the improvement of computational methods for studying neural networks will certainly bring invaluable insight in the field of perceptual processing, particularly in our understanding of the relationship between neural processing and behavior (Quian Quiroga and Panzeri, 2009), this present work highlights the fact that complex processing of higher-order acoustic features occurs at the level of the single neuron.

The next step that logically follows this observation is to assess to what extent the acoustic features of the sounds can predict the neural responses, in the context of this processing of higher-order information.

D. How are the acoustic parameters of the calls used for individual voice discrimination in an auditory scene?

This question is an interesting perspective for future work. Indeed, having discussed above about the fact that bottom-up processes can reach a level of complexity that has not been shown before in the context of auditory scene analysis, it would be insightful to investigate how the spectro-temporal features of degraded sounds are processed by the highly discriminant neurons described in **manuscript 3**. More specifically, in future work I plan to compare the parameters that were found to be highly individualized in the acoustic analysis presented in **manuscript 1** to the spectro-temporal receptive fields of these neurons. Do the highly discriminant neurons use all the individualized parameters that were found at different distances? In other words, how much of the redundancy in the information coding of the calls is used by these neurons? Also, do the neurons extract the "best acoustic parameters" for individual discrimination, which were found to be different between short and long distance? Or does the neural code rely on a set of parameters that are averaged across distances, similar to the "all distances" template presented in **manuscript 1**?

Another interesting research axis will be to investigate the coding properties of ensembles of neurons and assess how their synchronization properties can generate an additional layer of auditory scene processing. Previous studies have indicated that temporal coherence between neural responses to stimuli could be critical to the processes of scene segmentation in the visual cortex (Engel et al., 1991) and auditory stream formation in the auditory cortex (Shamma et al., 2011). By investigating the coding properties of neural networks in our data obtained using multi-electrode arrays, we could find for example that combining the coding properties of neurons that show "short-distance coding templates" (i.e., high discrimination abilities at short and medium distance but not at long distance) with those of neurons showing "long-distance coding templates" (i.e., medium discrimination abilities at short and medium distance, and a discrimination performance at long distance that matches the best performance found in the acoustic analysis) could be useful in circumventing potential issues due for example to the fact that the most individualized acoustic parameters are not the same depending on distance. We could thus bring to light even more invariant responses to degraded calls than what we found studying the discrimination performance at the level of the single neuron. For example, I hypothesize that we could estimate the stimulus intensity from the ensemble responses; if this is verified, it would bring further insight on how the brain deals with this particular aspect of auditory scene analysis.

Comparing the biological answers to the issue of auditory scene analysis to results obtained from sound analyses and modeling may bring further insight into the role of bottom-up processes in the transition from the high fidelity encoding of the physical structure of sounds found in the lower auditory pathway to the higher-order interpretation of this physical structure found in the primary auditory cortex (Nelken and Ahissar, 2006). Furthermore, having found impressive discrimination abilities for a complex scene analysis task in single neurons, studying the interactions between these neurons, and the higher-order auditory representations permitted by ensemble coding, could help bring further insight in how the brain processes this seemingly easy and yet computationally challenging task that is scene analysis. Ultimately, this would help us come closer to making a link between the stimuli perceived, their processing in the brain and the subsequent behavioral response, and thus advance our understanding of the processes involved in animal communication.

REFERENCES

V. REFERENCES

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

Alain, C., Arnott, S. R. and Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *J. Exp. Psychol. Hum. Percept. Perform.* **27**, 1072–1089.

Allen, W. L. and Higham, J. P. (2013). Analyzing Visual Signals as Visual Scenes: Analyzing Visual Signals. *Am. J. Primatol.* **75**, 664–682.

Appeltants, D., Gentner, T. Q., Hulse, S. H., Balthazart, J. and Ball, G. F. (2005). The effect of auditory distractors on song discrimination in male canaries (*Serinus canaria*). *Behav. Processes* **69**, 331–341.

Aubin, T. and Jouventin, P. (1998). Cocktail–party effect in king penguin colonies. *Proc. R. Soc. Lond. B Biol. Sci.* **265**, 1665–1673.

Aubin, T. and Jouventin, P. (2002a). How to vocally identify kin in a crowd: The penguin model. In *Advances in the Study of Behavior*, pp. 243–277. Elsevier.

Aubin, T. and Jouventin, P. (2002b). Localisation of an acoustic signal in a noisy environment: the display call of the king penguin *Aptenodytes patagonicus*. *J. Exp. Biol.* **205**, 3793–3798.

Aubin, T., Jouventin, P., Hildebrand, C., Aubin, T., Jouventin, P. and Hildebrand, C. (2000). Penguins use the two–voice system to recognize each other. *Proc. R. Soc. Lond. B Biol. Sci.* **267**, 1081–1087.

Bailey, D. J., Rosebush, J. C. and Wade, J. (2002). The hippocampus and caudomedial neostriatum show selective responsiveness to conspecific song in the female zebra finch. *J. Neurobiol.* **52**, 43–51.

Bee, M. A. (2008). Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. *Anim. Behav.* **75**, 1781–1791.

Bee, M. A. and Klump, G. M. (2004). Primitive Auditory Stream Segregation: A Neurophysiological Study in the Songbird Forebrain. *J. Neurophysiol.* **92**, 1088–1104.

Billimoria, C. P., Kraus, B. J., Narayan, R., Maddox, R. K. and Sen, K. (2008). Invariance and Sensitivity to Intensity in Neural Discrimination of Natural Sounds. *J. Neurosci.* **28**, 6304–6308.

Blumstein, D. T. and Turner, A. C. (2005). Can the acoustic adaptation hypothesis predict the structure of Australian birdsong? *Acta Ethologica* **8**, 35–44.

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot Int.* **5**, 341–345.

Bolhuis, J. J. and Gahr, M. (2006). Neural mechanisms of birdsong memory. *Nat. Rev. Neurosci.* **7**, 347–357.

Boncoraglio, G. and Saino, N. (2007). Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. *Funct. Ecol.* **21**, 134–142.

Boumans, T., Vignal, C., Smolders, A., Sijbers, J., Verhoye, M., Van Audekerke, J., Mathevon, N. and Van der Linden, A. (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.

Bradbury, J. W. and Vehrencamp, S. L. (1998). *Principles of animal communication.* Sinauer Associates. Sunderland, MA.

Branstetter, B. K. and Finneran, J. J. (2008). Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*). *J. Acoust. Soc. Am.* **124**, 625.

Bregman, A. S. (1993). Auditory scene analysis: hearing in complex environments. In *Thinking in sound: the cognitive psychology of human audition* (ed. McAdams, S. E. and Bigand, E. E.), Oxford.

Bregman, A. S. and Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *J. Exp. Psychol.* **89**, 244.

Brémond, J. C. and Aubin, T. (1992). The role of amplitude modulation in distress-call. *Ethol. Ecol. Evol.* **4**, 187–191.

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

Brown, T. J. and Handford, P. (2000). Sound design for vocalizations: quality in the woods, consistency in the fields. *The Condor* **102**, 81–92.

Brown, T. J. and Handford, P. (2002). Why birds sing at dawn: the role of consistent song transmission. *Ibis* **145**, 120–129.

Brumm, H. and Naguib, M. (2009). Environmental acoustics and the evolution of bird song. In *Advances in the Study of Behavior* (ed. Naguib, M., Janik, V. M., Zuberbühler, K., and Clayton, N. S.), pp. 1–33. New York.

Brumm, H., Voss, K., Köllmer, I. and Todt, D. (2004). Acoustic communication in noise: regulation of call characteristics in a New World monkey. *J. Exp. Biol.* **207**, 443–448.

Burley, N. T., Parker, P. G. and Lundy, K. (1996). Sexual selection and extrapair fertilization in a socially monogamous passerine, the zebra finch (*Taeniopygia gullata*). *Behav. Ecol.* **7**, 218–226.

Butterfield, P. A. (1970). The pair bond in the zebra finch. In *Social behaviour in birds and mammals: Essays on the social ethology of animals and man*, pp. 249–278. London: J. H. Crook.

Charrier, I., Mathevon, N. and Jouventin, P. (2003). Vocal signature recognition of mothers by fur seal pups. *Anim. Behav.* **65**, 543–550.

Charrier, I., Pitcher, B. J. and Harcourt, R. G. (2009). Vocal recognition of mothers by Australian sea lion pups: individual signature and environmental constraints. *Anim. Behav.* **78**, 1127–1134.

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

Cosens, S. E. and Falls, J. B. (1984). A comparison of sound propagation and song frequency in temperate marsh and grassland habitats. *Behav. Ecol. Sociobiol.* **15**, 161–170.

Cynx, J., Lewis, R., Tavel, B. and Tse, H. (1998). Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata*. *Anim. Behav.* **56**, 107–113.

Dabelsteen, T. and Mathevon, N. (2002). Why do songbirds sing intensively at dawn? A test of the acoustic transmission hypothesis. *Acta Ethologica* **4**, 65–72.

Dabelsteen, T., Larsen, O. N. and Pedersen, S. B. (1993). Habitat-induced degradation of sound signals: Quantifying the effects of communication sounds and bird location on blur ratio, excess attenuation, and signal-to-noise ratio in blackbird song. *J. Acoust. Soc. Am.* **93**, 2206.

Dahmen, J. C. and King, A. J. (2007). Learning to hear: plasticity of auditory cortical processing. *Curr. Opin. Neurobiol.* **17**, 456–464.

Dent, M. L., McClaine, E. M., Best, V., Ozmeral, E., Narayan, R., Gallun, F. J., Sen, K. and Shinn-Cunningham, B. G. (2009). Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *J. Comp. Psychol.* **123**, 357–367.

Dooling, R. J. (1986). Perception of vocal signals by budgerigars. *Exp. Biol.* **45**, 195–281.

Doucet, S. M. (2003). Multiple sexual ornaments in satin bowerbirds: ultraviolet plumage and bowers signal different aspects of male quality. *Behav. Ecol.* **14**, 503–509.

Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631.

Duvall, D. (1979). Western fence lizard (*Sceloporus occidentalis*) chemical signals. I. Conspecific discriminations and release of a species-typical visual display. *J. Exp. Zool.* **210**, 321–325.

Eales, L. A. (1985). Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Anim. Behav.* **33**, 1293–1300.

Engel, A. K., König, P. and Singer, W. (1991). Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl. Acad. Sci.* **88**, 9136–9140.

Ey, E. and Fischer, J. (2009). The "acoustic adaptation hypothesis"—a review of the evidence from birds, anurans and mammals. *Bioacoustics* **19**, 21–48.

Fay, R. R. and Popper, A. N. (2000). Evolution of hearing in vertebrates: the inner ears and processing. *Hear. Res.* **149**, 1–10.

Foote, A. D., Osborne, R. W. and Hoelzel, A. R. (2004). Environment: Whale-call response to masking boat noise. *Nature* **428**, 910–910.

Forrest, T. G. (1994). From sender to receiver: propagation and environmental effects on acoustic signals. *Integr. Comp. Biol.* **34**, 644–654.

Forstmeier, W., Burger, C., Temnow, K. and Derégnaucourt, S. (2009). The genetic basis of zebra finch vocalizations. *Evolution* **63**, 2114–2130.

Fortune, E. S. and Margoliash, D. (1992). Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (*Taenopygia guttata*). *J. Comp. Neurol.* **325**, 388–404.

Fotheringham, J. (1997). Song transmission and auditory perception of distance in wood warblers (*Parulinae*). *Anim. Behav.* **53**, 1271–1285.

Fuller, R. A., Warren, P. H. and Gaston, K. J. (2007). Daytime noise predicts nocturnal singing in urban robins. *Biol. Lett.* **3**, 368–370.

Gentner, T. Q. (2004). Neural Systems for Individual Song Recognition in Adult Birds. *Ann. N. Y. Acad. Sci.* **1016**, 282–302.

Gentner, T. Q. and Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature* **424**, 669–674.

Gilbert, C. D., Sigman, M. and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron* **31**, 681–697.

Gish, S. L. and Morton, E. S. (2010). Structural adaptations to local habitat acoustics in Carolina wren songs. *Z. Für Tierpsychol.* **56**, 74–84.

Gold, J. I. and Watanabe, T. (2010). Perceptual learning. *Curr. Biol.* **20**, R46–R48.

Grace, J. A., Amin, N., Singh, N. C. and Theunissen, F. E. (2003). Selectivity for Conspecific Song in the zebra finch auditory forebrain. *J. Neurophysiol.* **89**, 472–487.

Hauber, M. E., Cassey, P., Woolley, S. M. N. and Theunissen, F. E. (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.

Hauser, M. D. (1996). *The Evolution of Communication*. MIT Press. London.

Hill, D. E. (1979). Orientation by jumping spiders of the genus Phidippus (*Araneae: Salticidae*) during the pursuit of prey. *Behav. Ecol. Sociobiol.* **5**, 301–322.

Holland, J. (1998). Degradation of wren Troglodytes troglodytes song: Implications for information transfer and ranging. *J. Acoust. Soc. Am.* **103**, 2154.

Holland, J., Dabelsteen, T., Bjørn, C. P. and Pedersen, S. B. (2001). The location of ranging cues in wren song: Evidence from calibrated interactive playback experiments. *Behaviour* **138**, 189–206.

Hsu, A., Woolley, S. M. N., Fremouw, T. E. and Theunissen, F. E. (2004). Modulation Power and Phase Spectrum of Natural Sounds Enhance Neural Encoding Performed by Single Auditory Neurons. *J. Neurosci.* **24**, 9201–9211.

Jouventin, P. (1982). Visual and vocal signals in penguins, their evolution and adaptive characters. *Fortschr. Verhaltensforsch.* **24**, 148.

Klump, G. M. (1996). Bird communication in the noisy world. In *Ecology and Evolution of Acoustic Communication in Birds* (ed. Kroodsma, D. E. and Miller, E. H.), pp. 321–338. Cornell University Press, Ithaca.

Klump, G. M. and Langemann, U. (1995). Comodulation masking release in a songbird. *Hear. Res.* **87**, 157–164.

Kondo, N. and Watanabe, S. (2009). Contact calls: information and social function. *Jpn. Psychol. Res.* **51**, 197–208.

Langmore, N. E. (1998). Functions of duet and solo songs of female birds. *Trends Ecol. Evol.* **13**, 136–140.

Larom, D., Garstang, M., Payne, K., Raspet, R. and Lindeque, M. (1997). The influence of surface atmospheric conditions on the range and area reached by animal vocalizations. *J. Exp. Biol.* **200**, 421–431.

Latinus, M., McAleer, P., Bestelmeyer, P. E. G. and Belin, P. (2013). Norm-Based Coding of Voice Identity in Human Auditory Cortex. *Curr. Biol.* **23**, 1075–1080.

Lengagne, T., Aubin, T., Jouventin, P. and Lauga, J. (1999). Acoustic communication in a king penguin colony: importance of bird location within the colony and of the body position of the listener. *Polar Biol.* **21**, 262–268.

Lewicki, M. S., Olshausen, B. A., Surlykke, A. and Moss, C. F. (2014). Scene analysis in the natural environment. *Front. Psychol.* **5**.

Lohr, B., Wright, T. F. and Dooling, R. J. (2003). Detection and discrimination of natural calls in masking noise by birds: estimating the active space of a signal. *Anim. Behav.* **65**, 763–777.

Luther, D. A. and Wiley, R. H. (2009). Production and perception of communicatory signals in a noisy environment. *Biol. Lett.* **5**, 183–187.

MacDougall-Shackleton, S. A., Hulse, S. H., Gentner, T. Q. and White, W. (1998). Auditory scene analysis by European starlings (*Sturnus vulgaris*): perceptual segregation of tone sequences. *J. Acoust. Soc. Am.* **103**, 3581–3587.

Maddox, R. K., Billimoria, C. P., Perrone, B. P., Shinn-Cunningham, B. G. and Sen, K. (2012). Competing Sound Sources Reveal Spatial Effects in Cortical Processing. *PLoS Biol.* **10**, e1001319.

Marler, P. and Slabbekoorn, H. (2004). *Nature's Music, the Science of Birdsong.* Elsevier Academic Press, London.

Marten, K. and Marler, P. (1977). Sound transmission and its significance for animal vocalization. *Behav. Ecol. Sociobiol.* **2**, 271–290.

Mathevon, N. and Aubin, T. (1997). Reaction to conspecific degraded song by the wren *Troglodytes troglodytes*: Territorial response and choice of song post. *Behav. Processes* **39**, 77–84.

Mathevon, N., Dabelsteen, T. and Blumenrath, S. H. (2005). Are high perches in the blackcap *Sylvia atricapilla* song or listening posts? A sound transmission study. *J. Acoust. Soc. Am.* **117**, 442.

Mathevon, N., Aubin, T., Vielliard, J., da Silva, M.-L., Sèbe, F. and Boscolo, D. (2008). Singing in the rain forest: how a tropical bird song transfers information. *PLoS ONE* **3**, e1580.

Mathevon, N., Koralek, A., Weldele, M., Glickman, S. E. and Theunissen, F. E. E. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of Crocuta crocuta. *BMC Ecol.* **10**, 9.

McCabe, S. L. and Denham, M. J. (1996). A Model of Auditory Streaming. *Adv. Neural Inf. Process. Syst.* 52–60.

Menardy, F., Touiki, K., Dutrieux, G., Bozon, B., Vignal, C., Mathevon, N. and Del Negro, C. (2012). Social experience affects neuronal responses to male calls in adult female zebra finches. *Eur. J. Neurosci.* **35**, 1322–1336.

Moore, R. C., Lee, T. and Theunissen, F. E. (2013). Noise-invariant Neurons in the Avian Auditory Cortex: Hearing the Song in Noise. *PLoS Comput. Biol.* **9**, e1002942.

Morton, E. S. (1975). Ecological sources of selection on avian sounds. *Am. Nat.* **109**, 17–34.

Moss, C. F. and Surlykke, A. (2001). Auditory scene analysis by echolocation in bats. *J. Acoust. Soc. Am.* **110**, 2207.

Mouterde, S. C., Duganzich, D. M., Molles, L. E., Helps, S., Helps, F. and Waas, J. R. (2012). Triumph displays inform eavesdropping little blue penguins of new dominance asymmetries. *Anim. Behav.* **83**, 605–611.

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

Nelken, I. and Ahissar, M. (2006). High-level and low-level processing in the auditory system: the role of primary auditory cortex. *Dyn. Speech Prod. Percept. P Divenyi Al Eds Amst. IOS Press* 343–353.

Nieder, A. and Klump, G. M. (1999). Adjustable frequency selectivity of auditory forebrain neurons recorded in a freely moving songbird via radiotelemetry. *Hear. Res.* **127**, 41–54.

Okanoya, K. and Dooling, R. J. (1987). Hearing in passerine and psittacine birds: a comparative study of absolute and masked auditory thresholds. *J. Comp. Psychol.* **101**, 7–15.

Park, T. J. and Dooling, R. J. (1986). Perception of degraded vocalizations by budgerigars (*Melopsittacus undulatus*). *Anim. Learn. Behav.* **14**, 359–364.

Patricelli, G. L. and Blickley, J. L. (2006). Avian communication in urban noise: causes and consequences of vocal adjustment. *The Auk* **123**, 639–649.

Perez, E. C., Elie, J. E., Soulage, C. O., Soula, H. A., Mathevon, N. and Vignal, C. (2012). The acoustic expression of stress in a songbird: Does corticosterone drive isolation-induced modifications of zebra finch calls? *Horm. Behav.* **61**, 573–581.

Pinaud, R. and Terleph, T. A. (2008). A songbird forebrain area potentially involved in auditory discrimination and memory formation. *J. Biosci.* **33**, 145–155.

Podos, J., Lahti, D. C. and Moseley, D. L. (2009). Vocal Performance and Sensorimotor Learning in Songbirds. In *Advances in the Study of Behavior*, pp. 159–195. Elsevier.

Pohl, N. U., Leadbeater, E., Slabbekoorn, H., Klump, G. M. and Langemann, U. (2012). Great tits in urban noise benefit from high frequencies in song detection and discrimination. *Anim. Behav.* **83**, 711–721.

Polley, D. B., Steinberg, E. E. and Merzenich, M. M. (2006). Perceptual Learning Directs Auditory Cortical Map Reorganization through Top-Down Influences. *J. Neurosci.* **26**, 4970–4982.

Pressnitzer, D., Meddis, R., Delahaye, R. and Winter, I. M. (2001). Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus. *J. Neurosci.* **21**, 6377–6386.

Quian Quiroga, R. and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* **10**, 173–185.

Rabinowitz, N. C., Willmore, B. D. B., King, A. J. and Schnupp, J. W. H. (2013). Constructing Noise-Invariant Representations of Sound in the Auditory Pathway. *PLoS Biol.* **11**, e1001710.

Ribeiro, S., Cecchi, G. A., Magnasco, M. O. and Mello, C. V. (1998). Toward a Song Code. *Neuron* **21**, 359–371.

Ritschard, M. and Brumm, H. (2012). Zebra finch song reflects current food availability. *Evol. Ecol.* **26**, 801–812.

Ryan, M. J. and Kime, N. M. (2003). Selection on long-distance acoustic signals. In *Acoustic Communication* (ed. Simmons, A. M., Popper, A. N., and Fay, R. R.), pp. 225–274. New York: Springer-Verlag.

Schneider, D. M. and Woolley, S. M. N. (2013). Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes. *Neuron* **79**, 141–152.

Schnitzler, H.-U. and Flieger, E. (1983). Detection of oscillating target movements by echolocation in the Greater Horseshoe bat. *J. Comp. Physiol. A* **153**, 385–391.

Seitz, A. R. and Dinse, H. R. (2007). A common framework for perceptual learning. *Curr. Opin. Neurobiol.* **17**, 148–153.

Shamma, S. A. and Micheyl, C. (2010). Behind the scenes of auditory perception. *Curr. Opin. Neurobiol.* **20**, 361–366.

Shamma, S. A., Elhilali, M. and Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* **34**, 114–123.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Information.* Illinois University Press. Urbana.

Simons, M. J. P. and Verhulst, S. (2011). Zebra finch females prefer males with redder bills independent of song rate--a meta-analysis. *Behav. Ecol.* **22**, 755–762.

Singh, N. C. and Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am.* **114**, 3394–3411.

Slabbekoorn, H. (2004). Singing in the wild: the ecology of birdsong. In *Nature's Music, the Science of Birdsong* (ed. Marler, P. and Slabbekoorn, H.), pp. 178–205. London.

Slabbekoorn, H. (2013). Songs of the city: noise-dependent spectral plasticity in the acoustic phenotype of urban birds. *Anim. Behav.* **85**, 1089–1099.

Snell-Rood, E. C. (2012). The effect of climate on acoustic signals: Does atmospheric sound absorption matter for bird song and bat echolocation? *J. Acoust. Soc. Am.* **131**, 1650.

Snyder, J. S. and Alain, C. (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychol. Bull.* **133**, 780–799.

Stripling, R., Milewski, L., Kruse, A. A. and Clayton, D. F. (2003). Rapidly learned song-discrimination without behavioral reinforcement in adult male zebra finches (< i> Taeniopygia guttata</i>). *Neurobiol. Learn. Mem.* **79**, 41–50.

Sueur, J., Aubin, T. and Simonis, C. (2008). Equipment review-Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics* **18**, 213–226.

Team, R. C. (2005). R: A language and environment for statistical computing.

Theunissen, F. E. and Shaevitz, S. S. (2006). Auditory processing of vocal sounds in birds. *Curr. Opin. Neurobiol.* **16**, 400–407.

Theunissen, F. E., Amin, N., Shaevitz, S. S., Woolley, S. M. N., Fremouw, T. and Hauber, M. E. (2004). Song Selectivity in the Song System and in the Auditory Forebrain. *Ann. N. Y. Acad. Sci.* **1016**, 222–245.

Van Staaden, M. and Römer, H. (1997). Sexual signalling in bladder grasshoppers: tactical design for maximizing calling range. *J. Exp. Biol.* **200**, 2597–2608.

Vates, E. G., Broome, B. M., Mello, C. V. and Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*Taenopygia guttata*). *J. Comp. Neurol.* **366**, 613–642.

Verhey, J. L., Pressnitzer, D. and Winter, I. M. (2003). The psychophysics and physiology of comodulation masking release. *Exp. Brain Res.* **153**, 405–417.

Vicario, D. S., Naqvi, N. H. and Raksin, J. N. (2001). Sex differences in discrimination of vocal communication signals in a songbird. *Anim. Behav.* **61**, 805–817.

Vignal, C., Mathevon, N. and Mottin, S. (2004a). Audience drives male songbird response to partner's voice. *Nature* **430**, 448–451.

Vignal, C., Attia, J., Mathevon, N. and Beauchaud, M. (2004b). Background noise does not modify song-induced genic activation in the bird brain. *Behav. Brain Res.* **153**, 241–248.

Vignal, C., Mathevon, N. and Mottin, S. (2008). Mate recognition by female zebra finch: analysis of individuality in male call and first investigations on female decoding process. *Behav. Processes* **77**, 191–198.

Von der Emde, G. (2004). Distance and shape: perception of the 3-dimensional world by weakly electric fish. *J. Physiol.-Paris* **98**, 67–80.

Watanabe, T., Nañez, J. E. and Sasaki, Y. (2001). Perceptual learning without perception. *Nature* **413**, 844–848.

Wilczynski, W., Ryan, M. J. and Brenowitz, E. A. (1989). The display of the blue-black grassquit: the acoustic advantage of getting high. *Ethology* **80**, 218–222.

Wiley, R. H. and Richards, D. G. (1982). Adaptations for acoustic communication in birds: sound transmission and signal detection. In *Acoustic Communication in Birds* (ed. Kroodsma, D. E. and Miller, E. H.), pp. 131–181. New York.

Wood, W. E., Yezerinac, S. M. and Dufty, J. (2006). Song sparrow (*Melospiza melodia*) song varies with urban noise. *The Auk* **123**, 650–659.

Woolley, S. M. N., Fremouw, T. E., Hsu, A. and Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat. Neurosci.* **8**, 1371–1379.

Yotsumoto, Y., Watanabe, T. and Sasaki, Y. (2008). Different Dynamics of Performance and Brain Activation in the Time Course of Perceptual Learning. *Neuron* **57**, 827–833.

Zann, R. (1984). Structural variation in the zebra finch distance call. *Z. Für Tierpsychol.* **66**, 328–345.

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

Zekveld, A. A., Heslenfeld, D. J., Festen, J. M. and Schoonhoven, R. (2006). Top–down and bottom–up processes in speech comprehension. *NeuroImage* **32**, 1826–1836.

SOLVEIG C. MOUTERDE

TITRE : DISCRIMINATION A LONGUE DISTANCE DES SIGNATURES VOCALES INDIVIDUELLES CHEZ UN OISEAU CHANTEUR : DES CONTRAINTES DE PROPAGATION AU SUBSTRAT NEURONAL.

L'un des plus grands défis posés par la communication est que l'information codée par l'émetteur est toujours modifiée avant d'atteindre le récepteur, et que celui-ci doit traiter cette information altérée afin de recouvrer le message. Ceci est particulièrement vrai pour la communication acoustique, où la transmission du son dans l'environnement est une source majeure de dégradation du signal, ce qui diminue l'intensité du signal relatif au bruit. La question de savoir comment les animaux transmettent l'information malgré ces conditions contraignantes a été l'objet de nombreuses études, portant soit sur l'émetteur soit sur le récepteur. Cependant, une recherche plus intégrée sur l'analyse de scènes auditives est nécessaire pour aborder cette tâche dans toute sa complexité. Le but de ma recherche était d'utiliser une approche transversale afin d'étudier comment les oiseaux s'adaptent aux contraintes de la communication à longue distance, en examinant le codage de l'information au niveau de l'émetteur, les dégradations du signal acoustiques dues à la propagation, et la discrimination de cette information dégradée par le récepteur, au niveau comportemental comme au niveau neuronal. J'ai basé mon travail sur l'idée de prendre en compte les problèmes réellement rencontrés par les animaux dans leur environnement naturel, et d'utiliser des stimuli reflétant la pertinence biologique des problèmes posés à ces animaux. J'ai choisi de me focaliser sur l'information d'identité individuelle contenue dans le cri de distance des diamants mandarins (*Taeniopygia guttata*) et d'examiner comment la signature vocale individuelle est codée, dégradée, puis discriminée et décodée, depuis l'émetteur jusqu'au récepteur.

Cette étude montre que la signature individuelle des diamants mandarins est très résistante à la propagation, et que les paramètres acoustiques les plus individualisés varient selon la distance considérée. En testant des femelles dans les expériences de conditionnement opérant, j'ai pu montrer que celles-ci sont expertes pour discriminer entre les signature vocales dégradées de deux mâles, et qu'elles peuvent s'améliorer en s'entraînant. Enfin, j'ai montré que cette capacité de discrimination impressionnante existe aussi au niveau neuronal : nous avons montré l'existence d'une population de neurones pouvant discriminer des voix individuelles à différent degrés de dégradation, sans entrainement préalable. Ce niveau de traitement évolué, dans le cortex auditif primaire, ouvre la voie à de nouvelles recherches, à l'interface entre le traitement neuronal de l'information et le comportement.

Mots-clés : communication acoustique, signature individuelle, dégradation due à la propagation, discrimination individuelle, analyse de scène auditive, oiseau chanteur.

TITLE: LONG-RANGE DISCRIMINATION OF INDIVIDUAL VOCAL SIGNATURES BY A SONGBIRD: FROM PROPAGATION CONSTRAINTS TO NEURAL SUBSTRATE.

In communication systems, one of the biggest challenges is that the information encoded by the emitter is always modified before reaching the receiver, who has to process this altered information in order to recover the intended message. In acoustic communication particularly, the transmission of sound through the environment is a major source of signal degradation, caused by attenuation, absorption and reflections, all of which lead to decreases in the signal relative to the background noise. How animals deal with the need for exchanging information in spite of constraining conditions has been the subject of many studies either at the emitter or at the receiver's levels. However, a more integrated research about auditory scene analysis has seldom been used, and is needed to address the complexity of this process. The goal of my research was to use a transversal approach to study how birds adapt to the constraints of long distance communication by investigating the information coding at the emitter's level, the propagation-induced degradation of the acoustic signal, and the discrimination of this degraded information by the receiver at both the behavioral and neural levels. Taking into account the everyday issues faced by animals in their natural environment, and using stimuli and paradigms that reflected the behavioral relevance of these challenges, has been the cornerstone of my approach. Focusing on the information about individual identity in the distance calls of zebra finches *Taeniopygia guttata*, I investigated how the individual vocal signature is encoded, degraded, and finally discriminated, from the emitter to the receiver.

This study shows that the individual signature of zebra finches is very resistant to propagation-induced degradation, and that the most individualized acoustic parameters vary depending on distance. Testing female birds in operant conditioning experiments, I showed that they are experts at discriminating between the degraded vocal signatures of two males, and that they can improve their ability substantially when they can train over increasing distances. Finally, I showed that this impressive discrimination ability also occurs at the neural level: we found a population of neurons in the avian auditory forebrain that discriminate individual voices with various degrees of propagation-induced degradation without prior familiarization or training. The finding of such a high-level auditory processing, in the primary auditory cortex, opens a new range of investigations, at the interface of neural processing and behavior.

Key words: acoustic communication, individual signature, propagation-induced degradation, individual discrimination, auditory scene analysis, songbird.

Discipline : Bioacoustique, Ethologie, Neurophysiologie.

Laboratoire : Equipe de Neuro-Ethologie Sensorielle, ENES-CNPS CNRS UMR 8195
Université Jean Monnet, 23 rue du Dr Paul Michelon, 42023 Saint-Etienne Cedex 2, France