

VOCAL PRODUCTION AND PERCEPTION
IN *STREPTOPELIA* DOVES

Beckers, Gabriël Joseph Leonhard
Vocal production and perception in *Streptopelia doves*
PhD dissertation Leiden University, Leiden 2003
Printed by: Optima Grafische Communicatie, Rotterdam
ISBN: 90-6734-036-7

Part of the research in this thesis was supported by a travel grant from the Netherlands Organization for Scientific Research (NWO).
The printing of colour graphs in this thesis was made possible by a grant from the Leids Universiteits Fonds.

Copyright © 2003 by G. J. L. Beckers. All rights reserved. Any individual may make a single copy of this thesis, or parts thereof, provided that (a) the copy is used solely for the purpose of private study or research and (b) the copy is made by the person requiring its use or by an employee of the institution to which he or she belongs, without charge beyond the actual cost of copying. Specific permission will not be required for multiple photocopy reproduction of this thesis to be used for bona fide educational purposes, provided this is done by a member of staff of the university, school or other comparable institution concerned for distribution without profit to student members of that institution, and provided that the copies are made from the original thesis. Copies of parts of this publication should always indicate its original source. For all other matters relating to the reproduction, storage or transmission of any part of this publication, by any means, electronic, mechanical, photocopying or otherwise, prior permission is required from the author.

VOCAL PRODUCTION AND PERCEPTION
IN *STREPTOPELIA* DOVES

PROEFSCHRIFT
TER VERKRIJGING VAN
DE GRAAD VAN DOCTOR
AAN DE UNIVERSITEIT LEIDEN,
OP GEZAG VAN DE RECTOR MAGNIFICUS DR. D. D. BREIMER,
HOGLERAAR IN DE FACULTEIT DER WISKUNDE EN
NATUURWETENSCHAPPEN EN DIE DER GENEESKUNDE,
VOLGENS BESLUIT VAN HET COLLEGE VOOR PROMOTIES
TE VERDEDIGEN OP DONDERDAG 9 JANUARI 2003
TE KLOKKE 14:15 UUR

DOOR
GABRIËL JOSEPH LEONHARD BECKERS
GEBOREN TE URMOND, THANS GEMEENTE STEIN
IN 1970

Promotiecommissie

Promotor: Prof. dr C. J. ten Cate
Referent: Prof. dr O. Næsbye Larsen (University of Southern Denmark)
Overige leden: Prof. dr J. J. Bolhuis (Universiteit Utrecht)
Prof. dr E. van der Meijden
Prof. dr ir J. L. van Leeuwen (Wageningen Universiteit)
Prof. dr G. A. Zweers

Aan Marjolein



Aan mijn Ouders

(Pulls off several layers at once.)

What an enormous number of swathings!
Isn't the kernel soon coming to light?

(Pulls the whole onion to pieces.)

I'm blest if it is! To the innermost centre,
it's nothing but swathings – each smaller and smaller...
Nature is witty!

[*Peer Gynt* — Henrik Ibsen, 1867]

Contents

chapter 1	General introduction and overview	11
chapter 2	Perceptual relevance of species differences in acoustic signal structure	19
chapter 3	Perceptual salience of acoustic differences between conspecific and allospecific vocalizations	35
chapter 4	Pure-tone birdsong by resonance filtering of harmonic overtones	53
chapter 5	Mechanisms of frequency and amplitude modulation in ring dove song	65
	Nederlandse samenvatting (Dutch summary)	85
	References	93
	Acknowledgements	103
	Curriculum vitae	105

1 General introduction and overview

1.1 The study of avian acoustic communication

Bird vocalizations are among the most prominent of all animal acoustic signals. This is partly so because birdsong is ubiquitous. Almost all bird species vocalize, and birds can spend considerable time singing. Another reason for the prominence of bird vocalizations is that they can be very complex. Not all birds sing complex songs, however. Some birds only produce a single unmodulated note, whereas others display a dazzling vocal virtuosity in long songs with many different notes. This illustrates a third interesting point: bird acoustic signals are also very diverse. Bird vocalizations are almost always distinctly different between species, and most species have multiple vocalization types, sometimes amounting to repertoires of over a hundred different song types. The abundance, complexity, and diversity of bird vocal communication make it an attractive subject for scientific research. In fact, birdsong has become a prime theme of investigation within the field of animal communication.

Besides fundamental scientific interest, there is another motivation for research on birdsong. The production of birdsong involves rapid and complex motor patterns, requiring coordinated neuromuscular activity of respiratory, vocal, and craniomandibular systems (Suthers et al., 1999). In this aspect, it resembles human speech. Moreover, in about half of the bird species learning is important for the development of normal vocalization behaviour. In addition to complex motor control, these birds' vocalization behaviours have numerous other important parallels with human speech (for a review see Doupe and Kuhl, 1999). Birdsong thus provides us with a unique model system for research into topics as motor control, and learning and memory, in a behavioural context that is similar to that of human speech.

As outlined by Tinbergen (1963), animal behaviours can be analysed using four different types of question. *(i)* What are their proximate, underlying mechanisms?, *(ii)* How do they develop during an animal's lifetime?, *(iii)* What are their functions?, and *(iv)* How did they evolve? Addressing one of these questions often requires one also to investigate some of the others. For

instance, if one is interested in the evolution of the elaborate songs in songbirds, it is very useful to know that a particular feature of songbirds is that they develop these songs by learning them from others.

The work described in this thesis makes part of a larger, ongoing research program that uses vocal communication in birds to understand the interaction of evolutionary processes and the mechanisms underlying behaviour. It uses vocal communication in *Streptopelia* doves as a model case, and addresses both questions concerning mechanism and questions concerning evolution and selection. Previous research in this program has been reported in the PhD dissertations of Slabbekoorn (1998), Ballintijn (1999), and de Kort (2002), and in scientific papers cited throughout this thesis. Those studies addressed such diverse topics as sound signal structure, function of sound parameters, behavioural ecology of communication, development of vocalizations, phylogeny of *Streptopelia* species, and morphology of the vocal apparatus. Largely lacking still are studies into perception of vocal signals and physiology of vocal production. Insight into these proximate mechanisms are interesting in their own right, but may also contribute to a better understanding of how communication mechanisms play a role in the evolution of bird vocalizations. Therefore, these two topics are the focus of my thesis.

Below I will give a brief introduction on vocal communication in *Streptopelia* doves. I will finish this chapter with an overview and the main conclusions of this thesis.

1.2 *Streptopelia* doves as a model system

The genus *Streptopelia*, family *Columbidae*, consists of 17 species with a known phylogenetic history (Johnson et al., 2001), most of which occur in Africa and South-East Asia, and some in Europe.

Streptopelia vocalizations are considered to be relatively simple and stereotyped (Slabbekoorn et al., 1999), and they develop normally without learning (Nottebohm and Nottebohm, 1971). They are generally categorized into calls and coos, based on the context of production and on acoustic structure (Baptista et al., 1997). In this thesis I will address coo vocalizations only. *Streptopelia* coos are divided in three different types: *perch-coos*, which are long-range signals used in male–male conflict and female attraction, *bow-coos*, which are short-range signals with similar functions, and *nest coos*, which are involved in nesting activities and pair bonding (Goodwin, 1983). In some species these coo types are acoustically different, whereas in others the acoustic differences are small or non-existent. Nevertheless, also in the latter case the distinction between coo types is usually maintained on the basis of dif-

ferences in accompanying behaviour, or the tempo in which series of coos (bouts) are produced.

Many *Streptopelia* species live in sympatry with other congeneric species, and some of these have a very similar appearance. Each species produces a species-specific perch-coo (Slabbekoorn et al., 1999), but short-distance coos (nest- and bow-coos) are not always distinctly different between species (unpublished data, de Kort and ten Cate, 2002). This suggests that *Streptopelia* doves use perch-coos also to identify conspecific birds. Experimental evidence from playback studies in the field shows that *Streptopelia* doves respond differently to conspecific and allospecific perch-coos (de Kort and ten Cate, 2001), so indeed do discriminate between these signals in the absence of visual cues.

Inter- and intraspecific differences in the acoustic structure of *Streptopelia* coo vocalizations are present in multiple features. One feature is customarily – although somewhat inappropriately – termed *temporal structure*. A coo's temporal structure is its pattern of on and off switching of phonation ('gating'). Parameters used to quantify this feature include the number and duration of individual sound elements, and of the silent intervals that separate these elements. Figures 1.1a-d show examples of differences in temporal structure of two species' perch-coos.

Another important feature differing between coos is *amplitude modulation structure*. In general, amplitude modulation is the variation of intensity of a continuous sound over time. A very distinctive form of amplitude modulation in *Streptopelia* doves is one that gives rise to a 'trill' quality. In this trill-type amplitude modulation, sound intensity is reduced to zero or almost zero in rapid cyclic sequences (see Fig. 1.1c, which shows a coo with an amplitude modulation of about 25 cycles per second). Often, sound is even momentarily interrupted, and a strict distinction between amplitude modulation structure and temporal structure does consequently not exist in these cases, at least not at the level of signal structure. Nevertheless, a distinction between the two features remains useful because of the different time scales in which they manifest themselves, and also because they are perceived as being of a different quality, at least by the human ear.

The third type of feature differing between coos is *frequency modulation structure*¹, which is the pattern of frequency variation over time. Differences occur in the patterns of time–frequency variation, but also in more general

¹It should be noted that the term 'frequency modulation' as I use it here, is in the sense in which it is generally used in the field of bioacoustics (Bradbury and Vehrencamp, 1998). The existing literature of *Streptopelia* vocal communication, however, uses this term strictly to indicate a special kind of frequency modulation, namely instanta-

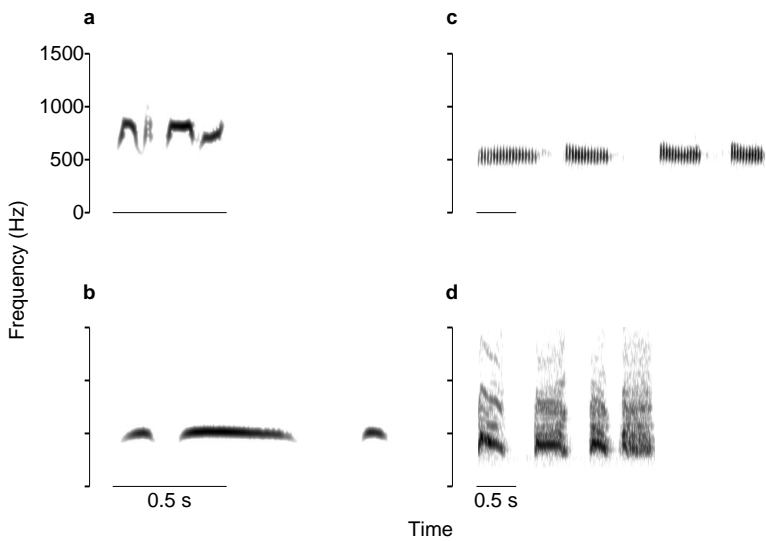


Figure 1.1: Spectrograms of perch-coos of (a) *S. vinacea*, (b) *S. decaocto*, (c) *S. turtur*, and (d) *S. orientalis*. Note that the time scales of subfigures are the same vertically, but not horizontally.

parameters, such as the ranges in which frequency is varied (e.g. Figs. 1.1a and b). Frequency modulation is often considered to be limited in nonsongbirds, mainly because frequency modulation in songbirds is, overall, more complex and extends over wider frequency ranges. Nevertheless, frequency modulation does occur in many nonsongbirds, and even limited frequency modulation has been shown to have communicative meaning in *S. decaocto* (Slabbekoorn and ten Cate, 1998; ten Cate et al., 2002).

A final class of features that may distinctly differ between coos concerns their *tonal quality*. Most dove coos are pure-tonal, which means that they only contain significant energy at a single frequency at any moment in time. Some dove coos, however, have energy at multiple discrete frequencies, or energy not restricted to discrete frequencies, but rather distributed in an irregular pattern over frequency bands (Fig. 1.1d). In the perception of humans at least, these differences give rise to distinctly different qualities of ‘tone’. Tonal quality of phonation is not necessarily uniform within coo vocalizations, and can change instantaneously even within continuous sound elements.

neous, discontinuous frequency shifts. In this thesis I will adhere to the original, and more widely used meaning, and refer to the latter with the term ‘frequency jumps’.

A main motivation of the work described in this thesis is to get a better understanding of the sound production mechanisms that underlie such differences in signal structure as described above, and to examine how relevant these features are, as perceived by doves.

1.3 Thesis overview

In this thesis I describe four studies into proximate mechanisms of vocal communication in *Streptopelia* doves.

In chapters 2 and 3, I examine which aspects of coo vocalizations are perceptually important to doves. In both studies I employed operant techniques in the laboratory, in which birds were trained with food rewards to respond to one category of sound stimulus, and not to respond to a second category. By monitoring the birds' behavioural response to probe stimuli – carefully chosen to share some characteristics with the first category and others with the second – it can be deduced which acoustic features the birds use to discriminate between the two categories.

In chapter 2, I examine the perceptual relevance of species-specific differences in perch-coo temporal structure and amplitude modulation structure. With acoustic analyses and statistical techniques, Slabbekoorn et al. (1999) showed earlier that within the *Streptopelia* genus, temporal ('note') structure is the most discriminative feature differing between species' perch-coos. However, acoustic signal characteristics that differ strongly between vocalizations of closely related species need not necessarily be the perceptually most important ones. Amplitude modulation ('trill') structure is also a prominent feature differing between species, and is more salient to the human ear (Beckers and ten Cate, unpublished results). We designed synthetic coo model stimuli based on natural coos of two species. These differed only in the two features of interest. Doves were trained to discriminate between these stimuli. By presenting hybrid probe sounds, varying one parameter at the time, we tested which parameters the animals used to make the discrimination. The results show that the majority of birds used both parameters to classify sound stimuli, although overall, amplitude modulation structure appeared somewhat more important.

The major advantage of the experimental design employed in the study of chapter 2 is that one can prove with certainty whether a specific acoustic feature is used for discrimination or not. On the other hand, it does not allow for the testing of more than two parameters at the same time, which makes testing the perceptual salience of larger number of parameters a tedious task. In addition, the use of artificial stimuli is a drawback since ultimately we are

interested in the perception of real vocalizations.

In **chapter 3**, we employed a method that is less direct in the identification of acoustic features relevant for the discrimination between species' perch-coos. However, it has the advantage that many parameters can be tested at the same time, and it allows the use of natural perch-coos instead of coo model stimuli. Doves were trained to respond to their conspecific coo only. The response strength to the perch-coos of the 12 other dove species was used as a measure for their perceptual similarity to conspecific perch-coos. The results show that the doves differentiate between allospecific coos: some are perceived as more similar to their own species' coo than others. With statistical techniques we then identified three acoustic features that correlate with perceptual similarity: coo duration, minimum frequency, and Wiener entropy. Their relative importance is determined by both their perceptibility and variance among coos. Duration is by far the most important feature in the discrimination between conspecific and allospecific vocalizations. Moreover, differences in duration are more salient to doves than expected on the basis of the threshold sensitivity of birds to changes in simple, short tones. This underlines that perceptual relevance of vocalization features should be studied using real vocalizations, or synthetic stimuli that resemble vocalizations in the parameters of interest.

Furthermore, a comparison of the results in chapter 3 with the phylogeny of the *Streptopelia* genus (Johnson et al., 2001) shows that there is no clear relation between the genetic distance among *Streptopelia* species and the perceptual similarity of their perch-coos. This is surprising, because cultural transmission does not play a role in the evolution of coo signal structure. All evolutionary change of the identified relevant features must therefore take place through natural and sexual selection, and genetic drift. Although such change is slower than it could be if cultural transmission were involved, it is apparently fast enough to obscure any clear phylogenetic pattern in the overall perceptual similarity of coos.

In the second half of this thesis, chapters 4 and 5, I report on two studies into sound production mechanisms of *Streptopelia* doves. The experimental techniques in the two experiments were identical: in spontaneously vocalizing birds we recorded air flow patterns through the trachea and air pressure patterns in different air sacs. The aims of the two studies, however, were very different.

With the study described in **chapter 4**, I addressed the mechanistic origin of pure-tonality in dove coos. Among the *Streptopelia* species' coos, most are pure-tonal, but some are not. There has been a long-standing discussion in the birdsong literature on whether or not pure-tone sound production re-

quires a source mechanism that is fundamentally different from a vibrating valve. Such a source is standard among tetrapods and usually gives rise to strong harmonic overtones in the sound. In pure-tone birdsong, however, harmonic overtones are lacking. We retrieved sound signals from the tracheal flow and air sac recordings, and show that the sound present inside the animal, close to the syringeal source, does contain harmonic overtones, which is consistent with a standard valve-type source. It is therefore unnecessary to invoke two separate, radically different sound generating mechanisms to explain the differences between pure-tone vocalizations and vocalizations with a more complex spectral structure. Apparently, sound is filtered to a pure tone before it radiates from the dove. This novel finding provides the most direct evidence for a source–filter mechanism in the production of pure-tone birdsong so far, and is also of interest beyond the context of birdsong. The existence of a multi-frequency, harmonic source sound that is filtered extends the already known parallels between human speech and birdsong to the level of acoustic production.

In **chapter 5**, I examine how ring dove vocalizations assume their specific modulation patterns in frequency and time domains. Air sac pressure and tracheal flow rate patterns reveal that trill-type amplitude modulation is the result of the cyclic opening and closure of a valve, instead of fluctuations in driving pressure as is the current explanation. Frequency modulation appears to be more complex than previously recognized, and consists of gradual, continuous time frequency patterns, punctuated by instantaneous frequency jumps. The gradual frequency modulation patterns correspond to pressure variation in the interclavicular air sac, which I argue is most likely the cause of such modulation. The cause of frequency jumps has not been identified, but can be explained on the basis of intrinsic properties of the vocal organ. Air sac pressure variation as a mechanism for frequency modulation is a novel addition to the ways birds regulate song frequency. It contrasts with the specialized syringeal musculature of songbirds, and may explain why song frequency in nonsongbird vocalizations is generally modulated within a limited frequency range.

Interestingly, the results in chapter 5 suggest that only two centrally controlled motor variables underlie the generation of modulation complexity in ring dove coos. Simple and smooth air sac pressure gestures determine the overall phonation and gradual frequency modulation patterns, while amplitude modulation is caused by the cyclic gating action of a valve. Much of the remaining complexity, such as oscillatory frequency modulation and frequency jumps, can be explained on the basis of mechanisms that require no central control: feedback and intrinsic nonlinear properties of the vocal or-

gan.

In **conclusion**, this thesis shows that despite the fact that *Streptopelia* species' perch-coos can be distinguished on the basis of temporal structure alone, doves use multiple features to discriminate between them. I have identified the overall most important acoustic features, which does not only provide us with a better insight into how dove coos are perceived by doves, but also indicates how vocal signals could diverge most efficiently in speciation events, from a perception point of view. The finding that differences in a temporal parameter are far more important to doves than they are to other birds, suggests that perceptual mechanisms have evolved in response to the fact that perch-coos predominantly differ in temporal structure. Even though this argument could also be turned around, it demonstrates that mechanisms and evolution of communication can be linked, and that Tinbergen's four question types should not be pursued in isolation.

Furthermore, in this thesis I show which peripheral mechanisms are involved in generating specific modulation patterns of coo vocalizations, and thus are responsible, on a proximate level, for many of the acoustic differences between species' coos. With the identification of air sac pressure as the causal factor in overall frequency modulation, I provide an explanation of why frequency modulation in doves, and perhaps in other nonsongbirds, is limited, as compared to songbirds. This mechanistic constraint in the modulation of frequency, in turn, provides an explanation of why species differences in dove coos are typically present in temporal parameters. The major weight that perceptual mechanisms attribute to a temporal parameter can therefore not only be linked to the evolution of species-distinguishing vocalizations, but also to their production mechanisms.

Taken together, this thesis reveals some of the biases, opportunities and constraints of the mechanisms underlying signals. In doing so, it provides more insight into how mechanisms of perception and production can form a shaping factor in the evolution of signals.

2 Perceptual relevance of species differences in acoustic signal structure

Summary Acoustic signal characteristics that differ strongly between vocalizations of closely related species need not necessarily be the perceptually most important ones. In the dove genus *Streptopelia*, temporal parameters have been identified as the most distinctive species-specific differences between perch-coo vocalizations. Using synthetic coo model stimuli and an operant design, we tested the perceptual relevance of species-specific differences in temporal structure and amplitude modulation structure for two partly sympatric species, *S. decaocto* and *S. chinensis*. The majority of birds used both parameters to classify sound stimuli, although overall, amplitude modulation structure appeared somewhat more important. In an additional experiment, we tested the same birds to examine whether they predominantly used the first element in the classification of coo model stimuli. Our results show that the characteristics of the full signal, rather than the first element, determine classification.

This chapter has also been published as: Beckers, G. J. L. and ten Cate, C. (2001). Perceptual relevance of species-specific differences in acoustic signal structure in *Streptopelia* doves. *Animal Behaviour*, 62: 519–525.

2.1 Introduction

Sympatric bird species often use dissimilar vocal signals, even when they are closely related and phenotypically similar. One explanation for this is that misidentification of conspecific signals because of confusion with those of closely related species can reduce the fitness of both sender and receiver. Such mistakes are a driving force for divergence of vocal signals, facilitating species discrimination. As Miller (1982) has pointed out, a scenario like this is likely to have most effect on long-range, unlearned sound signals, in particular when these are used in mate attraction and territorial advertisement.

Differentiation in acoustic signal structure can involve multiple parameters, the nature and magnitude of which will reflect constraints and biases in the mechanisms of signal production (sender) as well as signal perception (receiver). Although statistical techniques can be used to identify which parameters contribute most to species differences in signal structure, they do not reveal the role these parameters play in species discrimination. A parameter contributing considerably to the description of the difference in signal structure need not necessarily be equally important for the perception of signal differences. Psychophysical experiments have shown that proportionally identical changes in separate sound parameters may differ greatly in their perceptual salience. Birds, for instance, are about 10 times more sensitive to changes in frequency than in duration (Dooling, 1982). The integration of results from acoustic analysis and psychophysical studies, however, does not suffice to analyse which differences in various parameters between two acoustic signals are perceptually the most relevant ones. Differences in detection thresholds are usually measured in experiments using simple, arbitrary acoustic stimuli, such as short pure tones. For this reason, such experiments allow no direct conclusion about how a difference in one dimension of a more complex acoustic signal, for example in temporal structure, compares in perceptual salience to another dimension, for example frequency. Furthermore, the fact that a receiver can detect specific differences between sounds in a psychophysical experiment does not necessarily mean that it uses such differences to classify the sounds. To gain more insight into how animals perceive acoustic parameters in complex sounds such as vocalizations, we can use various psychoacoustic methods. Indeed a variety of studies have used such methods, ranging from the use of naturally occurring behavioural responses in amphibians (e.g. Gerhardt, 1991; Simmons and Bean, 2000) and insects (e.g. Wytenbach et al., 1996), to operant techniques in, for example, monkeys (e.g. LePrell and Moody, 1997; Sinnott and Saporita, 2000), songbirds (e.g. Genter and Hulse, 2000), nonsongbirds (e.g. Dooling et al., 1987a,b) and humans

(Iverson and Kuhl, 1995).

We used a psychoacoustic, operant technique to examine whether the prime parameter that characterizes the difference in vocalizations between related species is also the perceptually most salient one. The model system we used concerns the relative importance of two features that *Streptopelia* doves (turtle-doves) may use to distinguish their own perch-coo vocalizations from those of other species in the same genus, namely temporal structure and amplitude modulation structure. The genus *Streptopelia* consists of 17 species (Johnson et al., 2001), some of which look very similar, and almost all of which live in sympatry with other congeneric species. Each species produces its species-specific perch-coo (Slabbekoorn et al., 1999), which is a relatively simple, unlearned, long-range signal, used in male–male conflict and female attraction (Goodwin, 1983).

Temporal and amplitude modulation parameters vary conspicuously between these perch-coos, and playback experiments in the field have shown that territorial males of various *Streptopelia* species differentiate between conspecific and allospecific perch-coos (de Kort and ten Cate, 2001). An earlier study (Slabbekoorn et al., 1999) showed temporal structure to be the most distinctive feature differing between *Streptopelia* perch-coos. It has also been shown that *Streptopelia decaocto* is sensitive to the rhythm of perch-coos (Slabbekoorn and ten Cate, 1999). This may suggest that selection has operated on temporal parameters in particular and that these parameters are the most important ones for species discrimination. On the other hand, psychophysical studies suggest that differences in sound duration have to be relatively large in order to be noticed. So it can also be argued that species differences in this dimension are caused by evolutionary ‘drift’, which has occurred because the parameter is not perceptually relevant. The other feature, amplitude modulation structure, also differs between *Streptopelia* species, but not all species can be identified on this characteristic alone. Seven species produce perch-coos with cyclic amplitude modulation, which gives rise to a trill quality, at least to the human ear. Such vocalizations sound distinctly different from the perch-coos of the other species in the genus that lack cyclic amplitude modulation. However, the perceptual salience of this feature is unknown.

We compared the perceptual salience of interspecific differences in temporal and amplitude modulation structure. We created two synthetic sounds differing only in these parameters. These sounds resembled the species-specific perch-coos of two *Streptopelia* species, *S. decaocto* and *S. chinensis*. These species are sympatric in part of their geographical range (Baptista et al., 1997) and their perch-coos differ mainly in temporal structure and amplitude modulation structure (Fig. 2.1). We trained birds of both species to discriminate

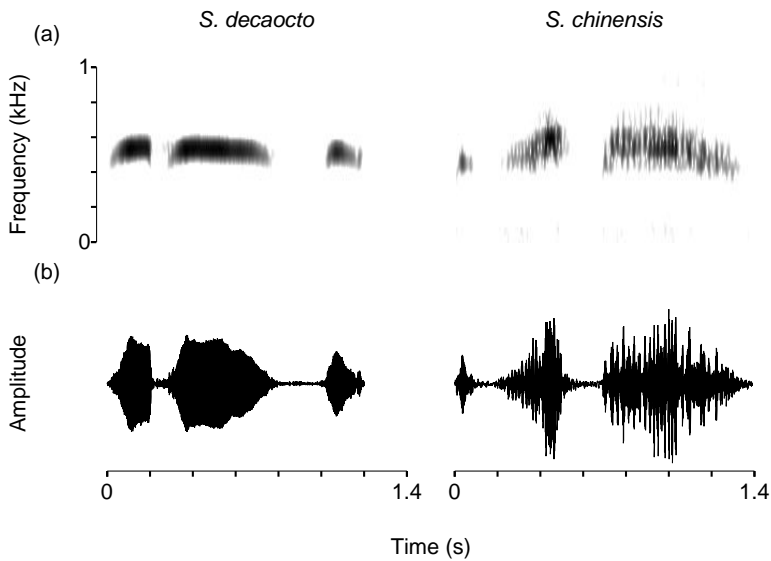


Figure 2.1: Perch-coos of *S. decaocto* and *S. chinensis*. (a) Spectrograms (calculated with a Short Time Fourier Transform, Hamming window, time resolution 15 ms, frequency resolution 87 Hz, dynamic range 25 dB) and (b) oscillograms.

between the sound stimuli, with a Go/No-go design (Hulse, 1995). In such a design, a bird is trained to respond differentially to, and thus classify, different complex acoustic stimuli. By presenting hybrid probe sounds, varying one parameter at the time, we can examine which parameters the animal uses to make the classification (Weary, 1990). This set-up allows precise control over the nature and degree to which parameters differ between the two signals. We can also measure differences in response strength, expressed as differences in number of operant responses, between species. We predicted that if species differences in temporal aspects of the perch-coo structure were due to selection on temporal structure, then this should also show up as the main parameter used for classification in our experiments. If, however, evolutionary drift caused differentiation in temporal structure, then we do not expect this parameter to be perceptually important. In this case, the amplitude modulation structure is expected to be the perceptually most salient one and used as a basis for classification of coo model stimuli in our experiments.

Different parts of a song can have different perceptual relevance (e.g. Johnsrude et al., 1994). For *S. decaocto*, field studies (Slabbekoorn and ten Cate, 1999) showed that the first element of the three-element perch-coo is the most important one in eliciting territorial responses. Without it, playback stimuli

did not evoke a significant response, whereas stimuli lacking the second or third element elicited normal responses. We carried out an additional experiment to test whether *Streptopelia* doves also classify coo model stimuli predominantly on the first element.

2.2 Methods

Subjects We used four adult *S. decaocto* (Eurasian collared doves; female number 45 and males number 15, 31, 47) and four adult *S. chinensis* (spotted doves; males number 138, 139, 152, 154) as subjects. The *S. decaocto* doves were originally captured from the wild in the Netherlands, and had been kept in our laboratory for more than 2 years. The *S. chinensis* doves originated from the wild in Indonesia, and had been kept for 1.5 years in our laboratory after we obtained them from a bird dealer in the Netherlands. All of the birds were naïve to operant experimental procedures. In spite of extensive training, two of the subjects, *S. chinensis* 152 and 154, did not acquire the operant task (see Results). Therefore, these birds could not be used for experimental tests.

On nonexperimental days the animals were housed in individual cages (80×60 cm and 60 cm high) in a room (maintained at ca. 22°C) containing other *Streptopelia* doves. Illumination was provided by high-frequency fluorescent lights, and the light:dark cycle was always 12:12 h. The diet of both species consisted of commercially available dry seed mix for turtle-doves, which was available ad libitum on nonexperimental days. On experimental days the same food was used as reinforcer. As subjects could initiate trials whenever they were motivated to do so, they could regulate their food supply according to their own needs. Every day we monitored the amount of food each bird had eaten, and, if necessary, a supplementary quantity was given. The birds were weighed twice a week. None of the birds showed weight loss below 90% of their free-feeding weight. After the experiments were completed, all subjects remained in our laboratory for other behavioural studies. We obtained permission for this study from the Animal Experiments Committee of Leiden University.

Operant test cages We used four identical operant test cages, each placed in a separate sound-attenuating chamber, for shaping, training and testing procedures. The test cages (60×50 cm and 60 cm high) were built of an open wooden frame, the top and three sides of which were covered with wire mesh. The fourth side was of plywood and held the control panel. On the top side, the control panel held a loudspeaker from which sound stimuli were presented. In the centre, two microswitch keys (diameter 1 cm) with built-in, red

light-emitting diodes (LEDs) were placed 16 cm apart. The left key functioned as the ‘observation key’, and the right as the ‘report key’. In between the keys a food dispenser was mounted. Keys and food dispenser were accessible to the bird via a perch. Illumination was provided by a high-frequency fluorescent light on top of the cage. We could observe the birds through one-way glass windows in the doors of the sound-attenuating chambers.

Each operant test cage was controlled by a small, custom-built computer device that was developed at the workshop of our institute. This device can be set up to function as a stand-alone machine, but we used it as an interface between the operant test cage and a personal computer. This allowed for faster and easier programming of more complex procedures. The device was also used to digitally store and play back sound stimuli.

Stimuli For baseline discrimination training, we designed two artificial sound stimuli, Train A and Train B (Fig. 2.2), which were modelled after the natural perch-coo vocalizations of *S. decaocto* and *S. chinensis* (Fig. 2.1), respectively. We constructed the sounds in such a way that they differed only in the two most important differences of the natural signals. The first difference concerned temporal structure. Both stimuli had three sound elements separated by two silent intervals, but the length of these structures, and thus the rhythm of the sounds, differed. The second difference concerned amplitude modulation structure. Train A had no amplitude modulation, while Train B was amplitude modulated. For details of sound parameters see Fig. 2.2. Signals were created digitally (10 kilosamples/s) with GoldWave software (version 4, GoldWave Inc., St. John’s, Newfoundland, Canada). Although our training stimuli were simplified versions of their natural counterparts and sounded artificial to human observers, the values chosen for temporal, amplitude modulation and frequency parameters fall within their natural ranges (Slabbekoorn et al., 1999).

Test sounds (Probes C, D, E and F) consisted of different combinations of the parameters of Train A and Train B (see Fig. 2.2).

Digital signals were transferred to the playback device with a 2821 Data Translation D/A board and a Frequency Devices 900 anti-imaging filter. The device subsequently resampled (26.8 kilosamples/s) the analogue signal by using its OKI MSM6388 sound processor. During the experiments, sound stimuli were generated by the same sound processor, which has a built-in filter to prevent imaging effects, and a Blaupunkt CB 4500 loudspeaker. All stimuli were matched for average sound level, which was 70 dB (A) SPL at the point where the bird’s head was during tests (as measured by a Cel-231 sound level meter, time weighting 125 ms).

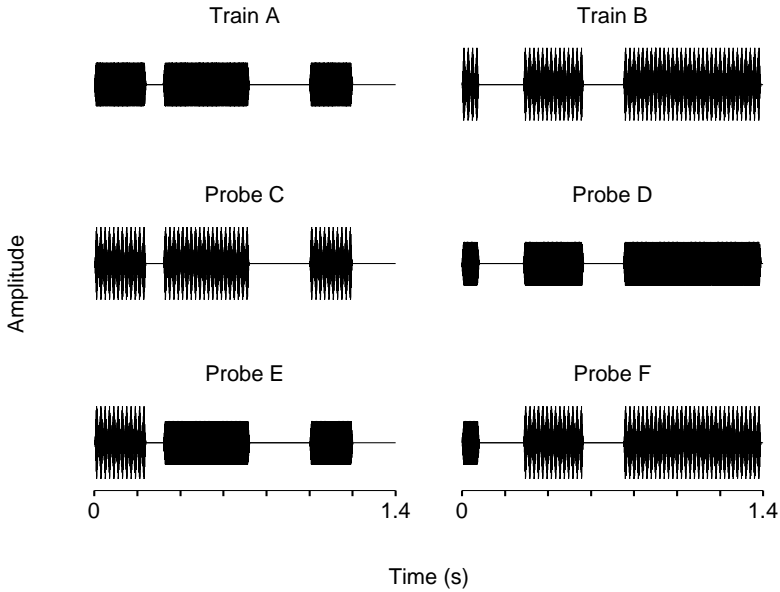


Figure 2.2: Oscillograms of sound stimuli. Train A and Train B were used for discrimination training, Probes C and D for tests in experiment 1, and Probes E and F for tests in experiment 2. If E_i designates the i th sound element and I_i the i th silent interval of a stimulus, then the temporal structure of Train A, Probe C and Probe E is E_1 240 ms, I_1 80 ms, E_2 400 ms, I_2 280 ms, E_3 200 ms, and that of Train B, Probe D and Probe F E_1 80 ms, I_1 200 ms, E_2 280 ms, I_2 180 ms, E_3 640 ms. Amplitude-modulated sound elements were created with a symmetric triangular modulating signal with a period of 20 ms. The carrier signal of all stimuli was a 600 Hz pure tone (sine wave), and all elements had linear rise and fall times of 15 ms.

Baseline discrimination training procedure To train birds to discriminate between the training stimuli Train A and Train B we used a Go/No-go procedure, in which they learned to peck a key in response to one stimulus and to withhold pecking in response to the other (Hulse, 1995).

Initially, the birds were autoshaped to peck the lit observation key for the presentation of a Go-stimulus, which was immediately followed by a food reward. In the next stage they had to peck the lit observation key to get a Go-stimulus, and when this sound had finished they had to peck the lit report key to get the food reward. When they performed this task well, key lights were gradually dimmed and finally turned off to prevent the LED light from overshadowing the auditory stimulus (Cynx and Clark, 1998).

After the shaping stage had finished, we started base-line discrimination training. A peck on the observation key initiated a trial with the random (without replacement) presentation of either a Go-stimulus or a No-go-stimulus ($P = 0.5$). Pecks at the report key in response to a Go-stimulus resulted in 3 s of access to food, and pecks in response to a No-go-stimulus resulted in 20 s time-out, during which the cage light was turned off. Pecks at either key during the presentation of a stimulus had no effect. If no peck on the report key had occurred within 2 s after the stimulus had finished, the trial ended and the bird could initiate a new trial by pecking the observation key. The assignment of Train A or Train B as Go-stimulus or No-go-stimulus was counterbalanced across birds of both species. However, the two *S. chinensis* that could not be trained to perform their complete operant task were in the same group. Therefore, counterbalancing was no longer possible for this species.

Training took place for 5 consecutive days a week, during which the birds remained in their operant test cages. Data of trials were collected every 24 h, and we defined each day as a separate ‘session’. For the remaining 2 resting days the birds were transferred to their home cage.

When the birds reached a level of 75% correct responses, reinforcement of trials was reduced to a random 90%, that is 10% of the correct responses were not followed by a reward and 10% of the incorrect responses were not followed by a punishment. We considered birds ready for testing if their discrimination level under this regime was 75% or more correct in three consecutive sessions.

Experiment 1: Use of temporal and amplitude modulation parameters Our objective in the first experiment was to determine which of the two acoustic features that differed between Train A and Train B, that is, temporal structure and amplitude modulation structure, the birds used for classification.

Test sessions were the same as training sessions, except that the proportion of unreinforced training trials was reduced to a random 5%, while in the remaining 5% Probe C or Probe D ($P = 0.5$) was given instead of the baseline training stimuli. Probe trials were never reinforced. Probe C had the temporal structure of Train A and the amplitude modulation structure of Train B, while Probe D had the temporal structure of Train B and the amplitude modulation structure of Train A (Fig. 2.2). By comparing the frequencies of Go-response of the probes to that of the training stimuli, we deduced which sound parameters the bird used. We finished a test when both probe stimuli had been given at least 50 times.

Experiment 2: Importance of first element After the birds had completed experiment 1, we used them for a second experiment. Our objective was to determine if the stimuli's first element alone had caused the pattern of probe classification obtained in experiment 1, or if the second and third elements also played a role. The procedure was the same as in experiment 1, except that instead of Probes C and D we used Probes E and F as test stimuli. The first elements of Probes E and F were identical to those of Probes C and D, respectively, while the second and third elements were identical to those of Trains A and B, respectively (Fig. 2.2). If the birds used only the first element to classify probe stimuli, then we expected their Go-response to Probes E and F of experiment 2 to be the same as their Go-response to Probes C and D of experiment 1, respectively.

Although all sound stimuli had been matched for root-mean-square sound pressure, we wanted to rule out the possibility that potential differences in perceived loudness had played a role in the discrimination of Trains A and B. Therefore, all birds underwent a post hoc test in which louder and softer versions of the training stimuli were used as probes (135 and 65% of root-mean-square sound pressure). If the birds generalized these probes to their corresponding training sounds, we concluded that sound level cues had not been important in the classification of stimuli.

Statistical analysis For each bird we pooled the data from all sessions within an experiment, to obtain the Go and No-go responses/stimulus. To assess which responses were significantly different from each other ($P < 0.05$), we applied a simultaneous test procedure (STP) based on G tests of independence (Sokal and Rohlf, 1995). This procedure identifies homogeneous subsets in a heterogeneous set of frequency data. On the basis of the outcome of the STP tests we assessed the relative importance of the acoustic parameters in

discrimination of the training stimuli for each bird qualitatively.

2.3 Results

Baseline discrimination training procedure Six of the eight subjects passed the criterion of 75% correct discrimination of the training stimuli Train A and Train B. The number of trials needed to acquire the criterion varied: 10250, 7215, 4748 and 6862 trials for *S. decaocto* 15, 31, 45, and 47, respectively, and 18308 and 7505 trials for *S. chinensis* 138 and 139, respectively. We failed to train the remaining two birds to perform their operant task: *S. chinensis* 152 did not get past the shaping stage (despite 20 weeks of training), whereas *S. chinensis* 154 did but failed to attain the criterion of 75% correct baseline discrimination (we stopped the procedure after 50 000 trials).

The post hoc test for potential influences of loudness in discrimination training showed that loudness was not an important cue in classifying stimuli for any of the birds. Go-response ratios of probes were never significantly different from those of their corresponding training sounds.

Experiment 1: Use of temporal and amplitude modulation parameters For each bird, Fig. 2.3 shows the ratios of Go-response to each stimulus, and homogeneous subsets of stimuli as identified by the STP tests, that is which stimuli do not have significantly different Go-responses.

From these data, we assessed which parameters each bird used to classify the probe stimuli, as follows. (1) If the data consisted of two homogeneous sets, each consisting of a probe and a training sound, as in dove 31 (Fig. 2.3, experiment 1), then this shows that the parameter type that differed between the two stimuli of a set had not been important for the dove in differentiating Train A from Train B. Therefore *S. decaocto* 31 used amplitude modulation structure, but not temporal structure. (2) If either one or both probes were not in a set with a training stimulus, then the bird used both parameters for classification. Therefore dove 45 (Fig. 2.3, experiment 1) used both amplitude modulation structure and temporal structure. This must necessarily be true because if only one of two parameters had been used, the probes would be perceptually equivalent to the training stimuli and they would consequently not be in different sets from the training stimuli. The relative importance of the parameters in this case was found as follows. (1) If the probes were together in a set without a training sound, as in doves 45, 47 and 139 (Fig. 2.3, experiment 1), then both parameters were considered equally important. (2) If the probes were not in a set with each other, but at least one of them was not in a set with a training sound either, as in dove 138 (Fig. 2.3, experiment 1),

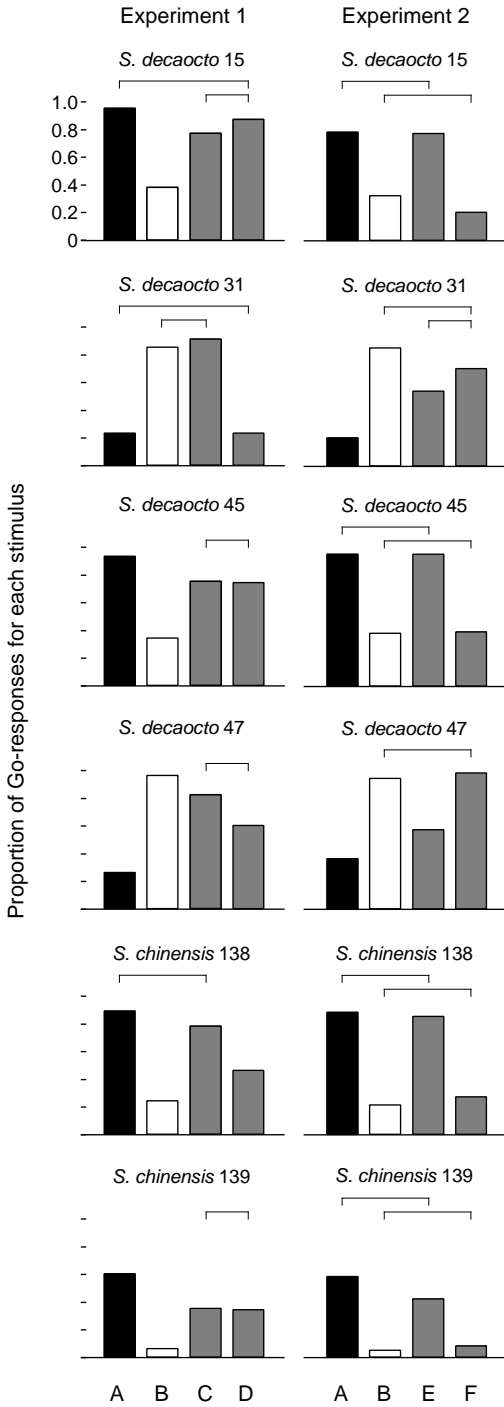


Figure 2.3: Classification patterns obtained in experiments 1 and 2. Letters beneath the bars correspond to the different stimuli (see Fig. 2.2). Horizontal brackets indicate which sound stimuli do not have significantly different Go-responses ($P < 0.05$).

or if they were both in a set but one of them was also in a set with a training sound, as in dove 15 (Fig. 2.3, experiment 1), then the relative importance of the two was found by looking at the relative position of the probes with respect to the training sounds.

The data show that five out of six birds used both parameters to classify the probe sounds. However, the birds varied in the relative weight of each parameter. None of the birds used temporal structure exclusively. One *S. chinensis* (138) gave more weight to temporal structure, but also used amplitude modulation structure. Three birds (*S. decaocto* 45 and 47 and *S. chinensis* 139) used both features, and did not give one feature significantly more weight than the other. One of these (*S. decaocto* 47), however, was close to giving amplitude modulation significantly more weight ($P = 0.08$). For *S. decaocto* 15 this difference is significant, so it used amplitude modulation, and to a lesser extent also temporal structure. One *S. decaocto* (31) exclusively used amplitude modulation.

Experiment 2: Importance of first element Figure 2.3 shows the ratios of Go-response to the stimuli of experiment 2, and the homogeneous subsets of stimuli as identified by the STP tests. Overall, most birds tended to classify Probes E and F with those training sounds with which they shared most features, that is Probe E with Train A and Probe F with Train B. Only in two birds (*S. decaocto* 31 and 47) was there a noticeable effect of difference between the first element of the probes and training stimuli. This is a first indication that classification is not based on features of the first element. Conclusive evidence, however, is provided by comparing the classification patterns with those found in experiment 1. If the subject had used the first element alone to classify the probe stimuli of experiment 1, then the Go-responses to Probes E and F would have to be the same as those to Probes C and D, respectively, since they would be perceptually equivalent. In fact, however, none of the subjects retained the classification pattern from experiment 1 to experiment 2. There was only one instance in which the change of probes between the experiments did not lead to a different classification: the responses of *S. chinensis* 138 to Probes C and E are both not significantly different from that to Train A. The same bird, however, did classify Probes D and F in a different way, indicating that it attended to the second and third elements.

2.4 Discussion

In experiment 1 we tested which acoustic features *S. decaocto* and *S. chinensis* use when they have to classify coo model stimuli that differ in two param-

ters only: temporal structure and amplitude modulation structure. One dove used amplitude modulation structure only. The other five used both features, but varied in the relative weights they attributed to temporal structure and amplitude modulation structure. The results show first that the *Streptopelia* species tested not only detected differences in the two parameters that constitute the most important species differences in their natural perch-coo vocalizations, but also used them in combination as a basis for sound classification. Second, the perceptual relevance of these parameters varied considerably between individuals, even within a species. This indicates that the relative contribution of temporal and amplitude modulation features to the classification process is not a fixed characteristic of individuals. Although the number of birds tested is too low to draw firm conclusions about interspecific differences, our data do show that potential species differences in perceptual relevance of the parameters tested are not discrete, that is there is overlap between the species in this characteristic.

In experiment 2 we tested whether the first element of a stimulus played a decisive role in the perceptual classification of coo model stimuli. Earlier studies suggested that for *S. decaocto* the first element of their perch-coo is the most important one in eliciting territorial responses in nature (Slabbekoorn and ten Cate, 1999). Our results show that attributing more weight to this part of the acoustic signal is not inherent to the working of perceptual classification mechanisms, as none of the birds predominantly used the first element to classify probe stimuli.

The average number of trials needed to reach baseline discrimination was 9148. Cynx (1995) trained domesticated pigeons, *Columba livia*, in on average 5250 trials for an auditory Go/No-go discrimination task. Slower discrimination training in our experiments could be due to a number of reasons. One is that the species tested are less tractable than those more typically used in psychoacoustic experiments. Also, our birds originated from the wild and were not hand reared. However, we tried to avoid possible adverse effects related to tractability by leaving the birds undisturbed in their operant cage for 5 days in a row during tests. In addition, all birds had been kept in our laboratory for more than 1.5 years prior to the experiments, and had adapted well to living in a cage. Another possibility explaining slow discrimination learning could be that the training stimuli are very difficult to discern for the subjects. However, this seems unlikely in our case. The differences in duration of elements and pauses of Train A and Train B are much greater than the known thresholds for this parameter in birds (Dooling, 1982) and from our results it is clear that the differences in amplitude modulation structure are at least as salient as those in temporal structure. We think that the most likely explanation for

slower discrimination training in our experiments is that this is inherent to the species used. Zebra finches, *Taeniopygia guttata*, trained by Cynx (1995) using the same experimental design and stimuli as for the pigeons, needed on average only 2250 trials. This shows that the number of trials needed very much depends on the species being trained.

To what extent can we relate our findings to vocal recognition in a natural context? We used synthetic sound models, which differ from natural perch-coo vocalizations in a number of aspects. A similar operant design with natural signals can address a number of interesting questions but it cannot reveal with certainty to which features of the vocalizations birds are listening (Dooling et al., 1992). For our purpose, we are necessarily restricted to a design with synthetic stimuli of which specific properties are manipulated while others are held constant. The use of sound models is justified as long as the information left out is unimportant for the question of interest. In our case, we reduced the set of varying parameters of natural coos by leaving out frequency modulation and components of amplitude structure that do not contribute to the trill quality. Both features are unlikely to be important when *S. decaocto* and *S. chinensis* have to discriminate their own species perch-coo from those of congeners. Their perch-coos show little frequency modulation (Fig. 2.1), and the frequency modulation that is present does not differ much between the species. Components of amplitude structure other than the one tested are not likely to be important for recognition in nature either, since they are aperiodic, slowly changing and relatively small, and therefore sensitive to unpredictable change through environmental influences, especially in such low-frequency, pure-tone signals (Michelsen, 1983). Temporal structure and amplitude modulation structure are probably also relevant features for the *Streptopelia* species not tested in our study, as these parameters constitute the main species differences in the genus. Other parameters that could be used are pitch and timbre. In general, however, both are likely to be of secondary importance because the majority of species (including the ones tested here) overlap in the carrier frequency they use, and produce relatively pure-tone vocalizations. So the two parameters we selected are likely to be important for perceptual classification in the wild.

The acoustic feature that has been identified as the most distinctive species difference between perch-coos in the *Streptopelia* genus, that is temporal structure (Slabbekoorn and ten Cate, 1999), appears not to be the perceptually most relevant one, although the majority of our birds used it. Limitations in how sound production mechanisms can change may constrain the evolution of amplitude modulation parameters more than that of temporal parameters in turtle-doves. Nevertheless, amplitude-modulated vocaliza-

tions do occur throughout the *Streptopelia* genus: seven of the 17 species have amplitude-modulated perch-coos (Slabbekoorn et al., 1999), while some of the remaining species produce other coo types that are amplitude modulated. Furthermore, potential species differences in the characteristics of amplitude modulation patterns, for example modulation rate and shape, have not yet been studied, but could play a role in species recognition or other signal functions. So perhaps amplitude modulation features in bird vocalizations have not received enough attention in comparative analyses as a feature contributing to interspecific differences. Lack of attention to amplitude modulation in acoustic analysis and its implications have been discussed recently for budgerigar, *Melopsittacus undulatus*, calls (Banta Lavanex, 1999).

Overall, we can conclude from our results that temporal structure, as suggested on the basis of acoustic and statistical analyses, is a factor that sympatric turtle-doves are likely to use for species discrimination in nature. The fact that another parameter, amplitude modulation, appears to be even more salient in perceptual classification illustrates that both signal variation and receiver sensitivity need to be studied to understand the importance of acoustic signal variability.

Acknowledgements We are grateful to Aldwin van Broekhoven, Renske Gilissen and Sebastiaan Ruinard who helped to carry out the experiments. We also thank Johan Bolhuis, Selvino de Kort, Jean Secondi, Nienke Terpstra and two anonymous referees for comments on the manuscript.

3 Perceptual salience of acoustic differences between conspecific and allospecific vocalizations

Summary During speciation, species-identifying vocal signals can acoustically diverge in many ways. Signal receivers have to be able to distinguish conspecific from allospecific vocalizations, and the perceptual salience of acoustic features is therefore expected to be an important factor in the evolution of such vocalizations. We tested how dissimilar the species-identifying perch-coos of 12 closely related turtle-dove species (genus *Streptopelia*) are, as perceived by one of its members, *S. roseogrisea*. With operant, psycho-acoustic methods we trained six doves to respond to their conspecific coo only. Responses to the perch-coos of the 12 other dove species were used as a measure for their perceptual similarity to conspecific perch-coos. Our results show that turtle-doves differentiate between allospecific coos: some are perceived as more similar to their own species' coo than others. With multiple-regression analysis we identified three acoustic features that correlate with these differences in perceptual similarity: coo duration, minimum frequency, and Wiener entropy. In contrast to findings in other bird species, duration is by far the most important feature in the discrimination between conspecific and allospecific vocalizations. The results suggest that this is not just due to duration being very different between coos of species, but also to a comparatively high perceptibility to differences in duration.

This chapter is also in press in Animal Behaviour as: Beckers, G. J. L., Goossens, B. M. A. and ten Cate, C. Perceptual salience of acoustic differences between conspecific and allospecific vocalizations in African collared-doves.

3.1 Introduction

Acoustic signals of closely related bird species are often distinctly different (Becker, 1982). This is especially true for long-distance vocalizations that are involved in mate attraction and territorial defense. Most likely, species-distinguishing characteristics of such signals are more important because the use of visual cues is difficult over larger distances, when sender and receiver are often out of view.

During speciation, vocal signals can acoustically diverge in many ways. Changes in e.g. song duration, element ('note') structure, frequency range, frequency modulation, amplitude modulation and tonality of vocalizations may be used to effectively separate the vocalizations of incipient species in signal space. Which parameters are more likely to change than others depends on several factors. One of these is that possibilities for change in the way sound signals are produced may favour particular directions for signal evolution, and disfavour others. Limits on the dynamics of vocal tract movements during song production, for example, probably constrain the development and evolution of syllable repetition rate in emberizid songbirds (Podos, 1997). Also, properties of the environment through which sound signals are propagated impose limitations on the acoustic characteristics of vocalizations, if they are to be transmitted efficiently. Higher frequency vocalizations, for example, do generally not carry as far in forested habitat as in open field (Wiley and Richards, 1982). But whatever the selection pressures or mechanisms of production, acoustic changes in vocalizations can only be selected for if they are salient to the receiver. Characteristics of perceptual systems are therefore also expected to play a role in signal evolution.

The perceptual salience of acoustic features in vocalizations can only be studied with psychoacoustic experiments in the laboratory, since playback experiments in the field do not allow one to distinguish the salience of features from their functional relevance. The role of acoustic features in discriminating vocalizations has been studied with psychoacoustic techniques in a number of bird species, e.g. budgerigars, *Melopsittacus undulatus*, (Brown et al., 1988; Dooling et al., 1987b, 1992), zebra finches, *Taeniopygia guttata*, (Okanoya and Dooling, 1991; Dooling et al., 1992), canaries, *Serinus canaria*, (Dooling et al., 1992), Bengalese finches, *Lonchura striata domestica*, (Okanoya and Kimura, 1993), great tits, *Parus major*, (Weary, 1990), red-winged blackbirds, *Agelaius phoeniceus*, (Sinnot, 1980), and brown-headed cowbirds, *Molothrus ater*, (Sinnot, 1980). None of these studies, however, address the perceptual differences in species-identifying vocalizations of closely related species, making it difficult to relate the results to an evolutionary context.

In this paper we examine the perceptual salience of acoustic differences in vocalizations of closely related species, by using turtle-doves (genus *Streptopelia*, 17 species, Johnson et al. (2001)) as a model system. Turtle-doves produce long-distance vocalizations, called perch-coos (see Fig. 3.1), that are highly stereotypic within species, but distinctly different between species. Slabbekoorn et al. (1999) showed that species-identity of individual birds can be correctly established on the basis of their perch-coo alone. Short-distance coo types (bow- and nest-coos), however, are not always distinctly different between species (de Kort and ten Cate, 2002). This suggests that turtle-doves use perch-coos to identify conspecific birds. Experimental evidence from playback studies in the field shows that turtle-doves respond differently to conspecific and allospecific perch-coos (de Kort and ten Cate, 2001), so indeed do discriminate between these signals in the absence of visual cues.

We reported earlier on the perceptual relevance of species-specific differences in perch-coos for two turtle-dove species, *S. decaocto* and *S. chinensis* (Beckers and ten Cate, 2001). Using an operant design and synthetic coo model stimuli differing only in the parameters of interest, we showed that both species use amplitude modulation structure and temporal structure for acoustic discrimination. The advantage of this experimental design is that we can prove with certainty whether a specific acoustic feature is used for discrimination or not. On the other hand, it does not allow for the testing of more than two parameters at the same time. *Streptopelia* species' perch-coos differ in several acoustic features, such as duration, temporal sound element structure, frequency range, peak frequency, frequency modulation pattern, amplitude modulation (trilled elements), and tonal structure (see Fig. 3.1 for examples). This makes testing the perceptual salience of all these parameters a tedious task. In addition, the required use of artificial stimuli is a drawback since ultimately we are interested the perception of real vocalizations.

In the current study, we therefore applied a different approach. We made use of the doves' differential behaviour to natural conspecific and allospecific perch-coos, as a method to investigate the perceptual salience of acoustic differences between species-specific signals. By employing psychoacoustic, operant methods we trained domesticated *S. roseogrisea* to peck a key in response to conspecific perch-coos only. In two methodologically different tests the number of key pecks in response to the perch-coos of 12 other dove species were used as a measure for their perceptual similarity to conspecific perch-coos. We then used the differential response scores of allospecific coos to identify underlying acoustic variables that correlate with the doves' perception of coo similarity.

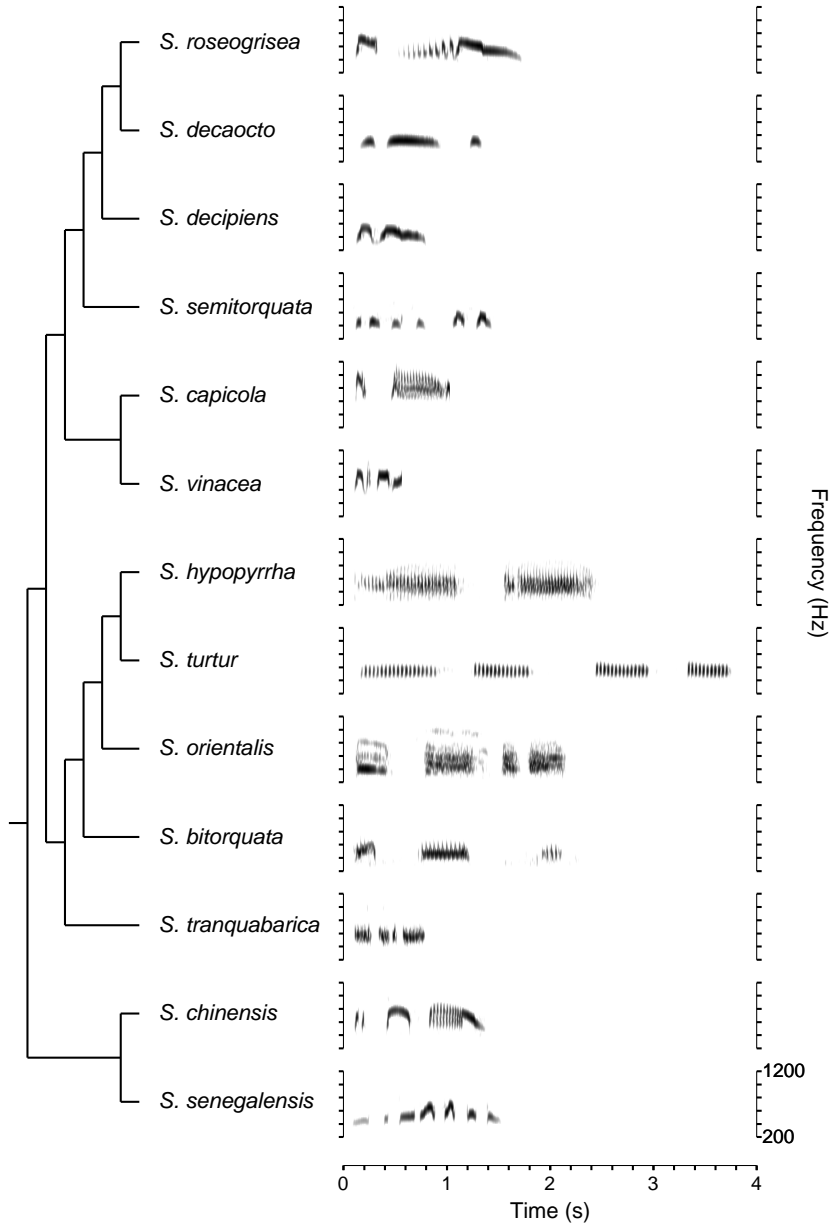


Figure 3.1: Phylogenetic relationship of 13 *Streptopelia* species and spectrograms of their perch-coos (Johnson et al., 2001). Branch lengths are not proportional to genetic distance.

3.2 Materials and Methods

Subjects We used one female and five male ring doves, a domesticated form of the African collared-dove *S. roseogrisea*. Although this domesticated form is sometimes also referred to as “*S. risoria*”, it is considered to be the same species (Baptista et al., 1997). All birds had been obtained commercially, and were experimentally naïve. When not participating in tests, animals were housed individually in cages (80 l × 60 w × 60 h cm), in a room that did not contain any other birds. Each day birds were transferred to an operant test cage for training or testing sessions. For half of the birds sessions lasted from 9 am to 5 pm, and for the other half from 5 pm to 9 am. Lights were switched on at 8 am and off at 8 pm, so the latter group effectively had a session that lasted for three hours in the afternoon and one hour in the morning. The birds’ diet consisted of a commercially available dry seed mix for turtle-doves, which they could normally only obtain by performing their operant task during the sessions. Birds could initiate trials whenever they were motivated to do so, and hence could regulate their own food supply. After each session we measured the amount of food each bird had eaten, and, if necessary, a supplementary quantity was given. Subjects were weighed twice a week to make sure their body weight did not fall below 90% of their free-feeding weight. After the experiments were finished, all subjects remained in our laboratory for other behavioural studies. This study was approved by the Animal Experiments Committee of Leiden University (UDEC nr. 98072).

Operant apparatus We used three identical operant cages, each placed in a separate sound-attenuating chamber, for training and testing procedures. The test cages measured 60 × 50 × 60 cm (l × w × h), and were built of an open wooden frame, the top and three sides of which were covered with wire mesh. The fourth side was of plywood and held the control panel. On the top side, the control panel held a loudspeaker from which sound stimuli were presented. In the centre, two microswitch keys (diameter 1 cm) with built-in, red light-emitting diodes (LEDs) were placed 16 cm apart. The left key functioned as the ‘observation key’, and the right as the ‘report key’. In between the keys a food dispenser was mounted. Keys and food dispenser were accessible to the bird via a perch. Illumination was provided by a high-frequency fluorescent light on top of the cage. We could observe the birds through one-way vision screens in the doors of the sound-attenuating chambers. All walls of the sound attenuating chambers (except the floor) were completely covered with sound absorbing foam (4 cm thick) to reduce adverse effects due to the reflection of sound waves.

Each operant test cage was controlled by a PC and custom-written software. A small, custom-built computer device that was developed at the workshop of our institute functioned as an interface between computer and operant test cage. This device was also used to digitally store and play back sound stimuli.

Stimuli We selected six perch-coo vocalizations from different individuals of each of 13 *Streptopelia* species (for examples see Fig. 3.1). One of these was *S. roseogrisea*, the species which the test subjects belong to.

We used field-recorded coos if they were available and had a good signal-to-noise ratio. These coos originated from different individuals at sites in Cameroon, Uganda, the Philippines, Great Britain, and the Netherlands. For five of the 78 coos we did not have suitable field recordings, so we used recordings from our lab (two *S. chinensis*, one *S. orientalis*, and two *S. tranquebarica*). Although the subjects in our study were domesticated *S. roseogrisea*, we used coo recordings from wild *S. roseogrisea*. We made this choice to avoid the introduction of systematic, recording-related differences between conspecific and allospecific coos in our tests. To the human ear, perch-coos of domesticated and wild *S. roseogrisea* do not sound differently, and visual inspection of spectrograms (personal observation) and acoustic analyses (Slabbekoorn et al., 1999) revealed no differences between the coos of the two forms.

Coos were sampled (44.1 kilosamples/s, 16 bit resolution) with a Card-Deluxe soundcard (Digital Audio Labs) and CoolEdit software (Syntrillium Software Corporation). Using the computer program Praat (version 3.9.11 for Windows, available from Paul Boersma and David Weenink, University of Amsterdam) we band-pass filtered all signals to remove sound energy outside the signal's frequency range, which was determined visually for each individual signal in a spectrogram (calculated with a Short Time Fourier Transform, Gaussian window, time resolution 15 ms, frequency resolution 87 Hz, dynamic range 40 dB). Filtering was applied in the frequency domain with Praat's built-in filtering function (Hann-like band shape, 50 Hz smoothing). After filtering, all stimuli were matched for root-mean-square sound amplitude. Digital signals were transferred to the playback device with the CardDeluxe soundcard. The device subsequently resampled (26.8 kilosamples/s) the signal using its OKI MSM6388 sound processor. During the experiments, sound stimuli were generated by the same sound processor, which has a built-in filter to prevent imaging effects, and a Blaupunkt CB 4500 loudspeaker. The playback level was set to 65 dB(A) SPL at the point where the bird's head was during tests, as measured by a Cel-231 sound level meter, time weighting 125 ms.

Shaping and baseline training We employed two different operant paradigms as methods to assess how much the different allospecific coos resembled conspecific coos perceptually. Both methods are based on a Go/No-go procedure (Hulse, 1995).

First, the birds were autoshaped to peck the lit observation key for the presentation of one of six conspecific perch-coos, selected at random without replacement by the operating software. The presentation of the coo stimulus was immediately followed by a food reward. In the next stage they had to peck the lit observation key to get a conspecific coo stimulus, and when this sound had finished they had to peck the lit report key to get the food reward. When they performed this task well, key lights were gradually dimmed and finally turned off to prevent the LED light from overshadowing the auditory stimulus (Cynx and Clark, 1998).

After the shaping stage we started baseline training. A peck on the observation key initiated a trial with the random ($P = 0.5$, without replacement) presentation of either a conspecific coo stimulus or a silent period lasting for the duration of a random conspecific coo stimulus. Pecks at the report key in response to a coo sound resulted in 3 s of access to food, but pecks in response to a silent period resulted in 20 s time-out, during which the cage light was turned off. Pecks at either key during the presentation of a coo sound or silent period had no effect. If no peck on the report key had occurred within 2 s after the coo sound or silent period had finished, the trial ended and the bird could initiate a new trial by pecking the observation key. When the birds reached a level of 75% correct responses, reinforcement of trials was reduced to a random 85%, that is 15% of the correct responses were not followed by a reward and 15% of the incorrect responses were not followed by a punishment. We considered birds ready for testing when they reached a level of 75% or more correct trials under this regime in three consecutive sessions.

Test 1 In the first test, trials were identical to those in training sessions, except that the proportion of unreinforced training trials was reduced to a random 5%, while in the remaining 10% a random one of 12 allospecific coos was given. These probe stimuli were never reinforced. The expectation was that birds would peck in response to probe stimuli more often if they were perceptually more similar to the conspecific coos, and less often if they were more dissimilar. We finished a test when the number of probe sounds that had been given amounted to an average of 100 per species, which took on average 35 days. After the tests had finished, we determined the ratio of peck responses to each coo stimulus per individual bird.

Test 2 Immediately after the first test, all birds were subjected to a second test, which was identical to the first one, except that probe stimuli (the 12 allospecific perch-coos) now became No-go-stimuli. A pecking response to any allospecific coo was punished, while pecking responses to conspecific coos remained rewarded. So in contrast to test 1, the birds now actively had to withhold pecking in response to an allospecific sound in order to avoid punishment. There were no probe trials and no unreinforced trials. We expected birds to learn this task faster or better for allospecific coos that were perceptually more dissimilar to their conspecific coos. Hence, as in test 1, the ratio of peck responses to each allospecific coo stimulus was considered a measure of its perceptual similarity to the conspecific coos. We had decided *a priori* to exclude trial sessions (days) in which the birds achieved less than 55% or more than 95% correct trials, since these would not be informative with respect to the relative differences between responses to different allospecific coos. We ended this experiment after 35 days and calculated the pecking response ratio for each stimulus. Over the whole testing period we obtained on average 276 trials per tested species' coo (46 per individual stimulus) per bird.

Transfer test The idea behind training the birds on six different conspecific coos instead of one was to minimize the possibility that the conspecific reference was, by accident, atypical for *S. roseo-grisea*. Nevertheless, we wanted to make sure that the birds indeed had formed a proper reference of conspecific coos, instead of references to potential oddities in the acoustic features of the six training coos. Therefore we subjected the birds to a one-day transfer test, immediately following upon the end of test 2. The transfer test was identical to the sessions in test 2, with the one difference that all six conspecific perch coos (Go-stimuli) had been replaced with six different conspecific perch-coos. We predicted that if the birds' reference during test 2 had been representative of conspecific coos, then the birds would maintain their normal discrimination level (conspecific versus allospecific) after this replacement. To assess whether or not discrimination levels were significantly different before and after replacement, we used loglinear analysis (Sokal and Rohlf, 1995) in SPSS for Windows, version 10.1, with the two sessions (before and after replacement), the six birds, the two stimulus categories (conspecific and allospecific coos), and the Go-response frequencies as factors. A loglinear model was built through backward elimination of factors and interactions that did not contribute significantly (P to remove was < 0.05). If in the resulting model the Go-response frequencies of the two sound categories were independent of session, we concluded that the replacement of the conspecific coos used during training by a new set of conspecific coos had not resulted in an important

change in the birds' discrimination behaviour.

Perceptual similarity scores Because of the design of our tests we cannot directly use the response difference between a particular allospecific coo and the average conspecific coo as a measure for their perceptual similarity. The overall Go-response levels, and also the level of discrimination vary from bird to bird, which has an effect on absolute response levels, and the differences in response levels between conspecific and allospecific coos.

Since we are interested in differences between perceptual similarity of allospecific coos, we standardized the Go-responses of the allospecific coos per individual bird with a z -transform, so that the mean response became zero and the standard deviation one. These scores are referred to as 'perceptual similarity scores'. A higher score means that a particular allospecific coo is more similar to the conspecific coos.

Identification of acoustic correlates For all coo stimuli used in the operant tests, we measured the acoustic parameters listed in Table 3.1.

Analyses were carried out with Praat except for the features *frequency modulation*, *spectral continuity* and *Wiener entropy*, which were obtained using the program Sound Analysis 2 (Tchernichovski et al., 2000). Spectrograms in Praat were calculated with a Short Time Fourier Transform with a Gaussian window, time resolution 15 ms, frequency resolution 87 Hz, and dynamic range 40 dB. Frequency spectra were calculated by a Fast Fourier Transform of the whole coo waveform, zero-padded up to the nearest radix-2 number. The rhythm of coo sounds were defined as their intensity envelopes (frame length 25 ms), and differences between the rhythm of coos was calculated by cross-correlation. The program Sound Analysis was used with its standard settings, except for the calculation of 'pitch', which was based on the peaks of power spectra.

The acoustic measurements of the six conspecific perch-coos were averaged. For each separate allospecific coo we then calculated the difference of each parameter with the conspecific average. These differences were in turn divided by the conspecific average. The advantage of expressing acoustic differences in fractions of the conspecific average value is that this makes it possible to quantify the comparison of different parameters (e.g. peak frequency and duration). This way, we can not only identify parameters that are used to discriminate between conspecific and allospecific perch-coos, but also determine their relative perceptual salience.

To examine which acoustic features of the allospecific coos can explain the differences in perceptual similarity scores, we employed multiple regres-

Acoustic feature	Description
Number of elements	Number of separate sound elements
Duration	Time span between the start of first sound element to the end of final sound element
Total duration elements	Sum of duration of separate sound elements
Dutycycle	Proportion of coo that consists of sound (= total duration elements / duration coo)
Rhythm	Intensity envelope over time
Minimum frequency	Lowest coo frequency visible in spectrogram
Maximum frequency	Highest coo frequency visible in spectrogram
Frequency span	Maximum frequency – minimum frequency
Peak frequency	Frequency with the highest amplitude in frequency spectrum
Spectrum bandwidth	Bandwidth of frequency spectrum 20 dB below peak frequency
Lowest frequency of band	Lowest frequency in spectrum at 20 dB below peak frequency
Highest frequency of band	Highest frequency in spectrum at 20 dB below peak frequency
Frequency modulation	Average slope of frequency change over time ^a
Spectral continuity	Measure of continuity of spectral features over time ^a
Wiener entropy	Measure of randomness of sound wave ^a
Fraction trill	Proportion of coo that is amplitude-modulated
Trill rate	Modulation frequency of amplitude-modulated (part of) coo

^aProcedures for calculating these parameters are given in Tchernichovski et al. (2000).

Table 3.1: Acoustic features measured from complete perch-coos. All frequency measures have also been determined separately for trilled parts of perch-coos. For details on measurements see text.

sion analysis (Sokal and Rohlf, 1995), using SPSS for Windows. One of the prerequisites of this type of analysis is that the independent variables entered into the model are not strongly correlated with each other. Therefore we first performed a Pearson correlation test on all combinations of parameters. If a parameter pair had a correlation coefficient of more than 0.7 we did not enter the parameter that had the lowest correlation with perceptual similarity into the regression analysis. Regression models were built with a stepwise method, with the criterion of $P < 0.05$ for entering, and $P > 0.1$ for the removal of variables. When necessary, we applied a natural logarithmic transform to variables to satisfy the assumptions of regression analysis

3.3 Results

All birds acquired their initial operant task of pecking the report key in response to conspecific coo stimuli and withholding pecks in the absence of such stimuli, at the criterion of 75% or more correct trials per session. With the introduction of allospecific probe coos in test 1, however, only four of the six birds showed significantly different pecking responses between the different allospecific coos (G test, $P < 0.05$). Overall, response ratios to allospecific probe coos appeared to be high: an average of 0.81 versus 0.90 for conspecific coos. When allospecific coos became No-go stimuli, in test 2, all birds acquired their new operant task. The Go-response ratio to allospecific coos was lower than in test 1: on average 0.67 versus 0.94 for conspecific coos.

Acoustic differentiation among allospecific coos Analysis of variance with similarity scores grouped per coo species showed that there are significant differences in perceptual similarity between coos of the 12 tested species (nested ANOVA, $F_{71,360} = 2.26$, $P < 0.001$, for test 1, and $F_{71,360} = 4.56$, $P < 0.001$, for test 2). Figure 3.2 shows the mean perceptual similarity scores grouped per coo species for both tests. A correlation test confirmed that the overall results from test 1 and 2 are very similar (Pearson correlation, $r = 0.96$, $P < 0.001$).

Acoustic correlates of differentiation Since the perceptual similarity scores of individual coos in test 2 were based on more operant trials per individual dove (46 versus 17 in test 1) and on more individual doves (six versus four in test 1), we used the similarity scores of test 2 as the dependent variable in the multiple regression analysis to identify acoustic correlates of perceptual similarity. We found three acoustic parameters with significant linear regression functions (Table 3.2): *duration*, *minimum frequency* and the natural logarithm of *Wiener entropy*. A regression model consisting of all three variables

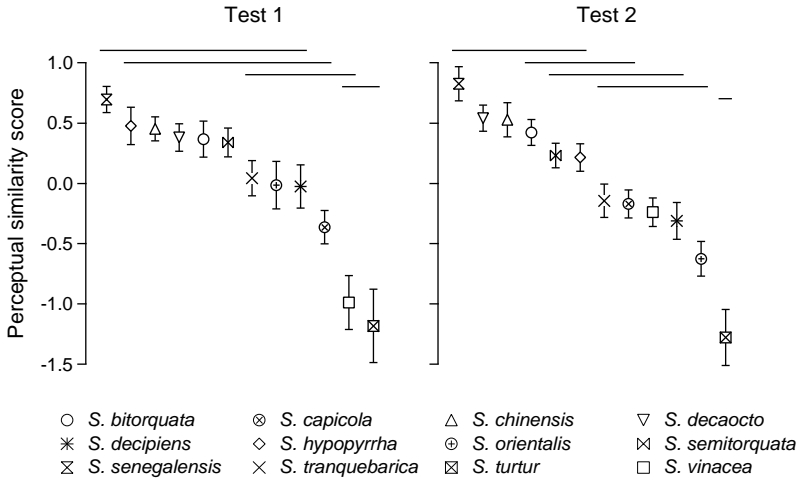


Figure 3.2: Perceptual similarity scores of allospecific perch-coos, grouped per species of origin. Symbols indicate the mean (\pm s.e.) per species' coo. Horizontal lines above the graphs indicate statistically homogeneous groups (nested ANOVA, Tukey's-b post hoc test, $\alpha = 0.05$).

is highly significant ($F_{3,68} = 81.8, P < 0.001$), and explains 78% of the variance in perceptual similarity scores (Fig. 3.3).

The acoustic features *duration* and *minimum frequency* had significant correlations higher than 0.7 with other parameters (*total duration elements*, and *peak frequency* and *lowest frequency of band*, respectively), which for this reason had not been entered in the regression analysis. Exchanging these parameters and rerunning the analysis resulted in significant models consisting of three parameters, but always with less variance explained.

Acoustic feature	R^2	$B (\pm \text{s.e.})$	P
Duration	0.548	-1.493 (± 0.112)	<0.001
Minimum frequency	0.154	-1.417 (± 0.226)	<0.001
Wiener entropy (ln)	0.081	-0.154 (± 0.031)	<0.001

Table 3.2: Multiple regression model of acoustic features on perceptual similarity. R^2 indicates the proportion of the variance of perceptual similarity that is explained by adding that particular feature to the model, and B is the partial regression coefficient. P values denote the significance of the partial regression equation.

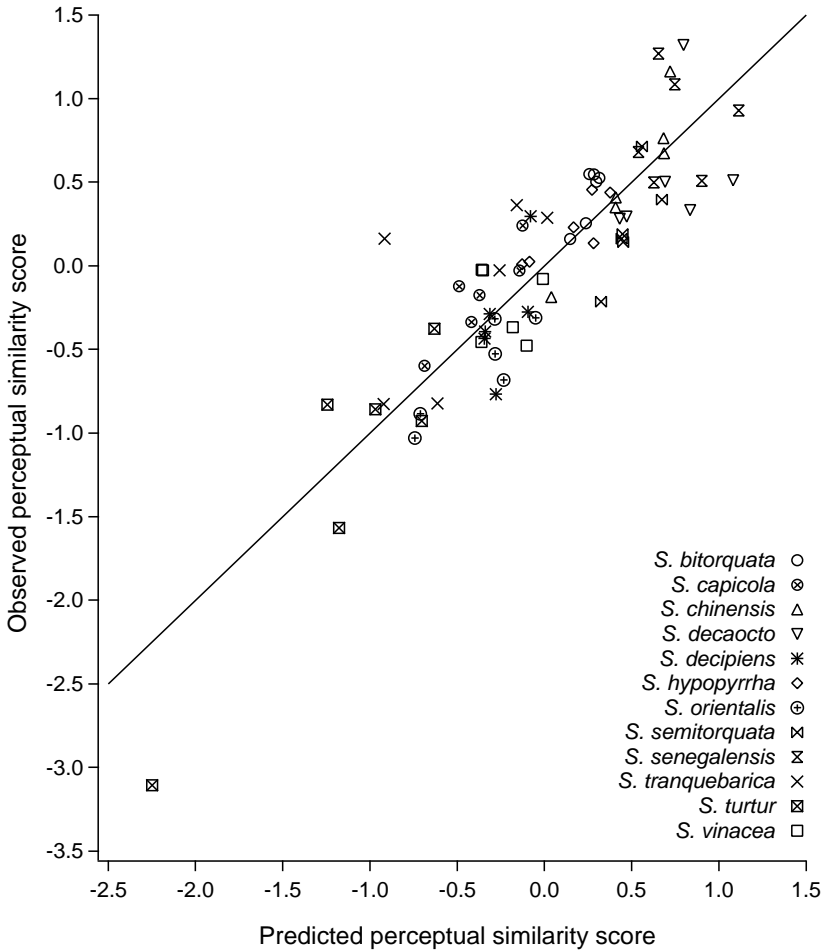


Figure 3.3: Predictive power of multiple regression model based on the acoustic parameters *duration*, *minimum frequency* and natural logarithm of *Wiener entropy*. The observed perceptual similarity values are standardized peck response ratios (z-transform), and the predicted perceptual similarity values result from the regression model (see Table 3.2). Symbols indicate the similarity score (mean of 6 subjects) of a particular perch-coo, where different symbols refer to the species of origin. The solid line indicates where predicted values equal observed values.

Transfer test The exchange of the six conspecific coos by six new ones in the transfer test after test 2 did not have a significant effect on discrimination levels (loglinear analysis, $\chi^2_5 = 1.27$, $P = 0.26$).

3.4 Discussion

Our results show that *S. roseogrisea* differentiates between allospecific perch-coos when discriminating their own species' perch-coo from that of others. We identified three acoustic features that correlate with this differentiation: *coo duration*, *minimum frequency* and *Wiener entropy*. The overall variance in perceptual similarity due to these features, expressed by the proportion of variance explained in the regression model (R^2), is determined by two factors: (1) their perceptibility, expressed by the magnitude of the regression coefficients (B), and (2) their acoustic variance among allospecific coos. So although differences in *coo duration* can explain 55% of the observed variance in perceptual similarity scores, and differences in *minimum frequency* only 15%, the perceptual salience of changes in either parameter are almost equal: the coefficients of their regression functions are -1.5 and -1.4 , respectively. The importance of temporal parameters such as duration in the perceptual differentiation of species' perch-coos fits with earlier observations of Slabbekoorn et al. (1999), who showed with acoustic and statistical analysis that temporal parameters are the most discriminative features between *Streptopelia* species' perch-coos. The small difference in perceptual salience of changes in a temporal and a frequency parameter, however, is a surprising finding. Just noticeable differences in duration of pure tones are known to be in the order of 10–20%, while this is about 1% for frequency (Dooling, 1982). Our results therefore confirm that it is important to distinguish between the perceptual salience of acoustic changes in complex vocalizations and underlying auditory abilities.

The fact that we identified three parameters with significant regression functions does not necessarily mean that those are used perceptually as such. It is possible that the birds' actual cues are correlated with the parameters that we measured. Moreover, it is possible that we missed relevant acoustic features, simply because we did not measure them, nor parameters that are correlated with them. On the other hand, the correlates we did find explain 78% of the observed variance, which leaves only limited room for as yet unidentified features.

The results from the transfer test show that the six conspecific perch-coos that we used in our experiments are representative for conspecific coos in general. Birds maintained their discrimination behaviour between conspecific

and allospecific coos after replacement of the conspecific coos used during training by new ones. It is likely that the subjects had formed a perceptual category of these conspecific coos, instead of using each of the six conspecific coos separately as a reference. A less likely explanation that cannot be excluded at present, is that the subjects could not tell the difference between the individual conspecific coos, since they are rather stereotypic. In any case, for our purposes it does not matter which of these two explanations is the correct one, because it is clear that whatever the birds referred to, it was not something that was specific for the selected training coos.

From the fact that duration turns out to be the most important parameter in distinguishing conspecific from allospecific perch-coos it follows that, overall, birds must have listened to the whole sound before deciding how to respond. However, in principle it could be possible that for very long coos the subjects did not wait for the stimuli to finish before responding. Even worse, they could have given up pecking before the stimulus was finished, with the result that in our experiments it would be classified as a No-go response. To check for such artifacts, we analyzed the response latency times for all sounds in both tests. Mean latency times for individual coos were always at least one standard deviation longer than the coo's duration, i.e. the birds waited for the stimulus to finish before responding. Besides the fact that this rules out the possibility of a response classification artifact, it also shows that birds have listened to completed coos before they respond, and are thus able to weight all parameters available.

The outcomes of tests 1 and 2 are very similar, although they originate from methodologically different experiments. This shows that our results are robust with respect to the type of operant task involved. On a species level, the only exception seems to be the coos of *S. vinacea*, which are perceived as more similar to conspecific coos in test 2 than in test 1 (Fig. 3.2). Given the fact that both methods gave similar results, we recommend a design similar to our second test (the discrimination design) for future research because of two practical advantages. First, the discrimination design generates many more trials per tested sound in a given amount of time, yielding more accurate perceptual similarity scores. Second, the difference between the overall response to conspecific coos and allospecific coos was larger in the discrimination design. Better separation of perceptual scores makes it a more powerful method to identify perceptual differences between vocalizations when they are relatively small.

Earlier we tested the perceptual relevance of temporal structure and amplitude-modulation structure ('trill') for *S. decaocto* and *S. chinensis* doves, when they have to discriminate between synthetic stimuli modelled after their

species' perch-coos (Beckers and ten Cate, 2001). The results showed that both features are used by the birds, with approximately similar weights. In the current study, amplitude-modulation structure did not show up as an important acoustic feature, but one aspect of temporal structure, namely its duration, did. The stimuli used in our previous experiment differed only in the two parameters of interest. Duration of the separate sound elements and the pauses between them differed markedly, but their total duration did so only slightly: 1.2 and 1.4 s. In our current tests, this would correspond to a difference in perceptual similarity score of less than 0.25. With only a small difference in duration available, the differences in amplitude-modulation apparently became significant cues for the subjects to perform their discrimination task. This illustrates an important point that should be made regarding the interpretation of our results: the fact that a particular parameter did not play a significant role in the discrimination tests does not mean that birds will not use that parameter when other cues are limited or not available. Overall, the three identified parameters are important when birds discriminate between conspecific and allospecific perch-coos. In specific cases, they may not be.

The major significance of coo duration in the discrimination of turtle-dove perch-coos contrasts sharply with the findings in other bird species. Duration does not play a significant role in the discrimination of conspecific from allospecific vocalizations in budgerigars, *Melopsittacus undulatus*, canaries, *Serinus canaria*, and zebra finches, *Taeniopygia guttata* (Dooling et al., 1992), and between allospecific vocalizations in starlings, *Sturnus vulgaris* (Dooling et al., 1992), and conspecific vocalizations in budgerigars (Brown et al., 1988). It is likely that at least part of the explanation for this lies in the fact that turtle-dove perch-coos differ distinctly in temporal structure (Slabbekoorn et al., 1999), of which duration is an important aspect. Nonetheless, the stimulus sets in the study of Dooling et al. (1992) did have significant differences in duration, and yet none of the four tested species used these as a cue for discrimination. This suggests that differences in duration are more salient to turtle-doves than they are to previously tested species. Such an increased sensitivity to temporal differences may have evolved in response to the fact that perch-coos differ predominantly in temporal features. Our methodological approach allows us to separate the effect of perceptibility of differences in acoustic features (regression coefficients) from that of their variance in the presented stimuli. The studies cited above, however, have focused on correlations between acoustic features and perceptual differences, and have not distinguished between feature perceptibility and variance. This makes a more quantitative comparison between species with respect to the perceptibility of differences in duration and other parameters not possible at present.

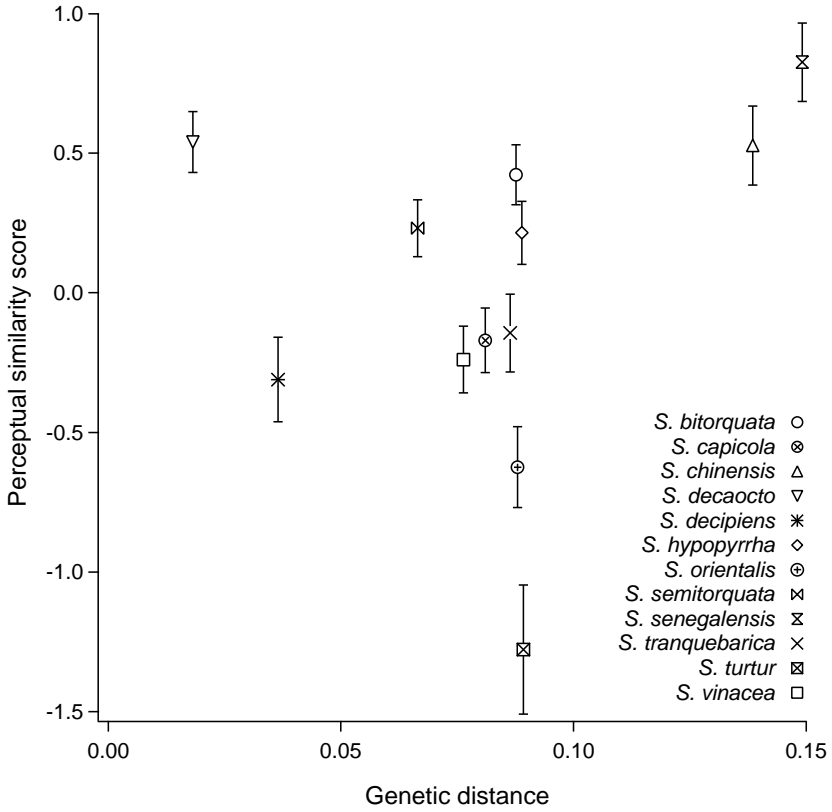


Figure 3.4: Genetic distance between *S. roseogrisea* and other species plotted against the perceptual similarity of their perch-coos. Genetic distances are in number of substitutions per site and originate from Johnson et al. (2001), and the perceptual similarity scores are from test 2 in the current study.

Recently, a well resolved phylogeny became available for almost all species in the genus *Streptopelia* (Johnson et al., 2001), which enables us to examine the perceptual similarity of vocalizations of the various species in a phylogenetic context. In Figure 3.4 we plotted the genetic distance between *S. roseogrisea* and the other species as reported by Johnson et al. (2001), against the perceptual similarity scores of their perch-coos as obtained in our study. This figure shows that the perceptually most similar coo is from the least related species, *S. senegalensis*, while at the same time the sister species of *S. roseogrisea*, *S. decaocto*, also has a high similarity score. Also from other comparisons it is noticeable that whether or not an allospecific vocalization is perceived as similar to the own species' coo is not apparent from its position in the phylogenetic tree. So despite the fact that turtle-doves do not learn their song (Nottebohm and Nottebohm, 1971), and hence an increased rate of evolutionary change of vocalizations through cultural transmission of mutations as in songbirds is not possible, evolutionary differentiation of perch-coos is fast enough to obscure a phylogenetic pattern in perceptual similarity. It seems likely that such rapid differentiation is partly due to the fact that the perceptually most salient acoustic feature, coo duration, can relatively easily be varied from a sound production point of view (Gaunt, 1986).

As a more general point this study demonstrates that the impact of various vocal parameters on signal recognition cannot be deduced directly from either between signal variation or psychophysical experiments concentrating on specific parameters. Establishing perceptual salience of various parameters in natural signals as done in the present study is thus an essential ingredient to interpret the perceptual relevance of species differences in signal structure.

Acknowledgements We would like to thank Selvino de Kort, Rob Lachlan, Albertina Leitão and Nienke Terpstra for commenting on the manuscript.

4 Pure-tone birdsong by resonance filtering of harmonic overtones

Summary Pure-tone song is a common and widespread phenomenon in birds. The mechanistic origin of this type of phonation has been the subject of long-standing discussion. Currently, there are three hypotheses. (i) A vibrating valve in the avian vocal organ, the syrinx, generates a multi-frequency harmonic source sound, which is filtered to a pure tone by a vocal tract filter ('source-filter' model, analogous to human speech production). (ii) Vocal tract resonances couple to a vibrating valve source, suppressing the normal production of harmonic overtones at this source ('soprano' model, analogous to human soprano singing). (iii) Pure-tone sound is produced as such by a sound generating mechanism that is fundamentally different from a vibrating valve. Here we present the first direct evidence for a source-filter mechanism in the production of pure-tone birdsong. Using tracheal thermistors and air sac pressure cannula's, we recorded sound signals close to the syringeal sound source during spontaneous, pure-tone vocalizations of two species of turtle dove. The results show that pure-tone dove vocalizations originate through filtering of a multi-frequency harmonic sound source.

This chapter is also submitted for publication elsewhere as: Beckers, G. J. L., Suthers, R. A. & ten Cate, C. Pure-tone birdsong by resonance filtering of harmonic overtones.

4.1 Introduction

There is a long history of discussion on how birds produce ‘whistled’ vocalizations (e.g. Cuvier and Duvernoy, 1846; Rüppell, 1933; Greenwalt, 1968; Casey and Gaunt, 1985; Fletcher, 1989; Gaunt and Nowicki, 1998). Normally, voiced sound production in tetrapods involves some kind of vibrating valve that modulates expiratory airflow in a pulsatile fashion (Stein, 1973). Sound generated this way is characterized by a multi-frequency harmonic spectrum, consisting of a fundamental frequency and a series of strong, harmonic overtones that are integer multiples of the fundamental. Pure-tone sound as found in whistled birdsong, however, is characterized by a simple harmonic frequency spectrum, in which essentially all acoustic energy is concentrated at a single frequency.

The apparent qualitative difference between whistled song and the sound that vibrating valves normally generate, has led to the suggestion that birds produce pure-tone signals in a fundamentally different way. The fact that the structure of the avian vocal organ, the syrinx, differs in many ways from the typical tetrapod vocal organ, the larynx, offers some support for this idea. Various mechanisms have been hypothesized to be the acoustic source in whistled song. Among these are a hole-tone whistle (Casey and Gaunt, 1985; Gaunt and Nowicki, 1998; Nottebohm, 1976), membranes vibrating in a string-like mode (Casey and Gaunt, 1985), and air-jet excited membrane oscillations (Fletcher, 1989). So far, only the involvement of an aerodynamic whistle has been subjected to experimental tests. The results, however, do not support this idea for the species investigated (Nowicki, 1987; Brittan-Powell et al., 1997; Ballintijn and ten Cate, 1998).

An alternative explanation for the absence of harmonic overtones in whistled birdsong is that a multi-frequency harmonic source sound is filtered to a pure tone by vocal tract resonances (Nowicki, 1987). The idea of a source-filter mechanism is derived from models of human speech production (Fant, 1960). If this model also proves correct in birds, then this would extend the already known parallels between human speech and birdsong (for a review see Doupe and Kuhl (1999)) to the level of acoustic production. Moreover, it would obviate the need to invoke a sound generator that is radically different from a vibrating valve. Nowicki (1987) provided experimental evidence to support the idea of a source-filter mechanism in birdsong by recording nine species of songbirds in a helium-enriched atmosphere. This light gas caused harmonic overtones to appear in vocalizations that were normally pure-tonal. Such overtones are expected when a vocal tract band-pass filter, normally centered around the fundamental frequency, shifts upward due to the increased

speed of sound.

The appearance of harmonic overtones in helium, however, could also be explained by assuming a different model for pure-tone sound production. In this alternative model, which is derived from human soprano singing, an overlap between vocal tract resonance and valve vibration frequencies results in a nonlinear feedback that suppresses production of harmonic overtones at the source (Gaunt and Nowicki, 1998; Nowicki and Marler, 1988; Hoese et al., 2000). In helium, the resonance frequency of the vocal tract shifts upward, breaking the coupling between vocal tract resonance and sound source, which results in the production of harmonic overtones.

Although the source–filter and soprano models differ in their prediction of whether or not the sound source produces harmonic overtones, both predict that it is the presence of specific vocal tract resonances that makes birdsong pure-tonal. Reports of helium studies in two other bird species, however, weaken the idea of vocal tract resonances as a general explanation for pure-tone birdsong. Brittan-Powell et al. (1997) concluded that in budgerigar contact calls the vocal tract filter only has a slight effect on the spectral content of the vocalization, which suggests that its narrow band frequency spectrum is produced as such by the syringeal source. Furthermore, Ballintijn and ten Cate (1998) concluded that the pure-tonality of collared dove coos is also due to syringeal source mechanisms, as coo vocalizations remain pure-tonal in helium.

Available evidence, therefore, does not allow one to distinguish between three hypotheses put forth to explain pure-tone birdsong production: (i) A vibrating valve produces a multi-frequency harmonic sound, which is filtered to a pure tone by a vocal tract resonance filter ('source–filter' model). (ii) Vocal tract resonances are coupled to the syrinx, causing a valve structure to vibrate sinusoidally and produce pure-tone sound ('soprano' model). (iii) A different, as yet unknown source mechanism produces pure-tone sound.

In the present study, we examined the first of these hypotheses, the source–filter mechanism, as an explanation for the origin of pure-tone coo vocalizations in two species of turtle-dove, *Streptopelia risoria* and *S. decaocto*. We recorded acoustic signals close to the birds' vocal organ, the syrinx, during normal, spontaneous vocalization. If pure-tonality is caused by filtering, then a multi-frequency source might reveal itself by the presence of harmonic overtones in its immediate vicinity. Vocal tract correlates of sound were recorded using micro-bead thermistors sensitive to air movement, implanted in the tracheal lumen just above the syrinx, and pressure sensitive transducers attached to cannula's inserted into the interclavicular air sac (ICAS), a cavity in which the syrinx is located. In addition, we recorded sound pressure in the cranial

thoracic air sac (CTAS), which is not in direct contact with the syrinx, but is connected to the ICAS and other air sacs through secondary bronchi. The results provide, to our knowledge, the first direct evidence that supports the hypothesis of a multi-frequency harmonic source sound in pure-tone bird-song.

4.2 Materials and Methods

Subjects We used four adult ring doves (*S. risoria*) and two adult Eurasian collared doves (*S. decaocto*) as subjects. Ring doves were obtained commercially, while the collared doves had originally been captured from the wild in the Netherlands, two years before the experiment.

Surgical procedures and recording of data The procedure to record air sac pressure and tracheal flow velocity signals is described in detail in Suthers et al. (1994), so we will only give a summary here.

Birds were anesthetized with isoflurane (Abbott Laboratories). A mid-line incision was made in the skin between the clavicles to expose the trachea as it entered the ICAS membrane. Air flow associated with near-field sound was measured with a microbead thermistor probe (Thermometrics, BB05JA202N) inserted into the tracheal lumen, just rostrally to the interclavicular membrane and about 1.7 cm (4 cm in two birds) cranial from the lateral tympaniform membranes of the syrinx. Thermistor wires were routed subcutaneously to connectors on a backpack that the birds wore. The flow velocity in the trachea was measured by a feedback circuit in which the current needed to maintain the heated thermistor at a constant temperature was proportional to the rate of air flow (Hector Engineering). Fluctuations in air pressure associated with sound production were monitored by a piezoresistive silicone diaphragm pressure transducer (Fujikura FPM-02PG) attached to an air sac cannula consisting of a flexible silastic tube (Dow Corning, ID = 1.02 mm, wall thickness = 0.57 mm). A cannula 18.5 cm long was inserted into the ICAS through a small hole in the interclavicular membrane from which it was routed subcutaneously to a backpack carrying the pressure transducer. The CTAS was cannulated by a similar tube 14 cm long inserted into the air sac through the abdominal wall just posterior to the last rib and a few mm lateral to the ventral midline. In each case the cannula extended 13 mm into the air sac and tissue adhesive was used to insure an air-tight seal around the cannula.

Vocalizations were recorded on a condenser microphone (Audio-technica AT 835b or Sennheiser MKH 40) placed 0.5–1 m in front of the cage. All signals

(emitted vocalization on microphone, ICAS pressure or tracheal flow velocity, and CTAS pressure) were recorded digitally (20 kilosamples/s) on a rotary storage recorder (Metrum Information Storage, model RSR 512) or on a DAT data recorder (TEAC, model RD135T). The low-frequency components, related to respiratory ventilation or phonatory motor patterns, of signals in the trachea and air sacs were removed with an analogue 100 Hz high-pass filter (Krohn-Hite, model 3550, or Princeton Applied Research Corporation, model 113) before recording, or by digital filtering after recording (see 'Data analyses'). We transferred the recorded signals from tape to a microcomputer by resampling (20 kilosamples/s), using a Data Translation DT-2821G board and a TTE 187 anti-aliasing filter (high cut-off at 8 kHz, stopband attenuation 60 dB per 1/3 octave).

Data analysis We considered recordings suitable for analysis if they had a good signal-to-noise-ratio and did not contain any recording artifacts, such as clipping. Tracheal sound signals were analyzed only if they were recorded no later than one week after thermistor implantation, since mucus and other material gradually accumulates on the thermistor, degrading its time constant and lowering the high frequency cut-off. Signals were analyzed using the software program Praat (available from Paul Boersma and David Weenink, www.praat.org), version 3.9.14 for Linux. All signals were filtered with a digital high-pass filter at 200 Hz, using a built-in function of the Praat program (frequency domain filter, Hann-like shaped band, 100 Hz smoothing).

We visually screened all signals for the occurrence of harmonic overtones using printed spectrograms, calculated by a short-time Fourier transform (frame length 30 ms, Gaussian window, time step 1 ms, frequency step 10 Hz, 40 dB dynamic range). In a selection of maximally 20 vocalizations for each combination of recorded variables, we then quantified the degree of puretonality of signals by measuring the intensity of the second ($2f_0$) and third ($3f_0$) harmonics, relative to that of the fundamental frequency (f_0). In puretone signals, the intensity of $2f_0$ and $3f_0$ should be very low. In the case of thermistor signals, we selected the first 20 recordings, since recording quality deteriorates over time. In the case of pressure signals, we selected 20 recordings on the basis of recording quality. First, we determined the dominant frequency in the power spectrum of the emitted coo sound. Using the spectrogram (see settings above), we then selected a portion of the coo where this frequency occurred with little frequency modulation. We calculated the mean fundamental frequency of a 200 ms time frame around this time point, using Praat's autocorrelation algorithm (Boersma, 1993). The 200 ms time frame of the emitted coo and concurrent time frames of the available ICAS, CTAS or tra-

cheal signals were then transformed to the frequency domain (8192 point FFT). The resulting spectra were divided up into 20 Hz frequency bins. The amplitudes of f_0 , $2f_0$ and $3f_0$ were determined by measuring the bin with the highest average amplitude in an 1/8 octave band around the expected harmonic frequencies, based on the (independent) calculation of mean fundamental frequency in the time domain.

Microbead thermistor probes and to a lesser degree also piezoresistive pressure transducers have non-linear response characteristics, which cause artificial introduction of some signal energy at harmonic overtone frequencies. We determined the extent of this effect by analyzing pure tones produced by a speaker and recorded with a microphone, thermistor, and pressure transducer (using the same procedure and analysis methods as described above). The results showed that within the frequency range of our dove coos ($500 \text{ Hz} < f_0 < 700 \text{ Hz}$) the intensity level of $2f_0$ and $3f_0$ introduced this way is about -24 and -33 dB, respectively, for thermistors, and lower than -40 dB (both $2f_0$ and $3f_0$) for pressure transducers (all dB values relative to the intensity of f_0). Similar calibration tests showed that errors in the measurement of $2f_0$ and $3f_0$ intensities due to the non-linear frequency response of the piezoresistive pressure transducers are less than 1 and 3 dB, respectively, for the frequency range of interest.

4.3 Results

Table 4.1 gives an overview of which signal combinations were obtained for each bird, as well as the number of recorded coos. Visual inspection of the spectrograms of these signals (Fig. 4.1) shows that both ICAS and tracheal sound signals always have a multi-frequency harmonic spectrum, with harmonic overtones at both even and odd multiples of the fundamental frequency. In low-noise ICAS recordings, one can even see the harmonic series reach the 11th harmonic, if the spectrogram's dynamic range is increased appropriately. Concurrent recordings of the emitted coos and CTAS signals never have a multi-frequency spectrum, but only consist of the fundamental frequency of the tracheal and ICAS spectra (Fig. 4.1).

A quantitative comparison of the intensity differences between f_0 and $2f_0$ or $3f_0$ confirms our qualitative observations. The intensities of $2f_0$ and $3f_0$, relative to that of f_0 , are much higher in the trachea and ICAS than in the emitted coo (Fig. 4.2). The intensity of $2f_0$ and $3f_0$ in CTAS is always below noise level (i.e., $< -30 \pm 4$ dB, mean \pm s.e., or -29 ± 4 dB at the expected frequencies of $2f_0$ and $3f_0$, respectively). CTAS signals can therefore be regarded as being pure-tonal.

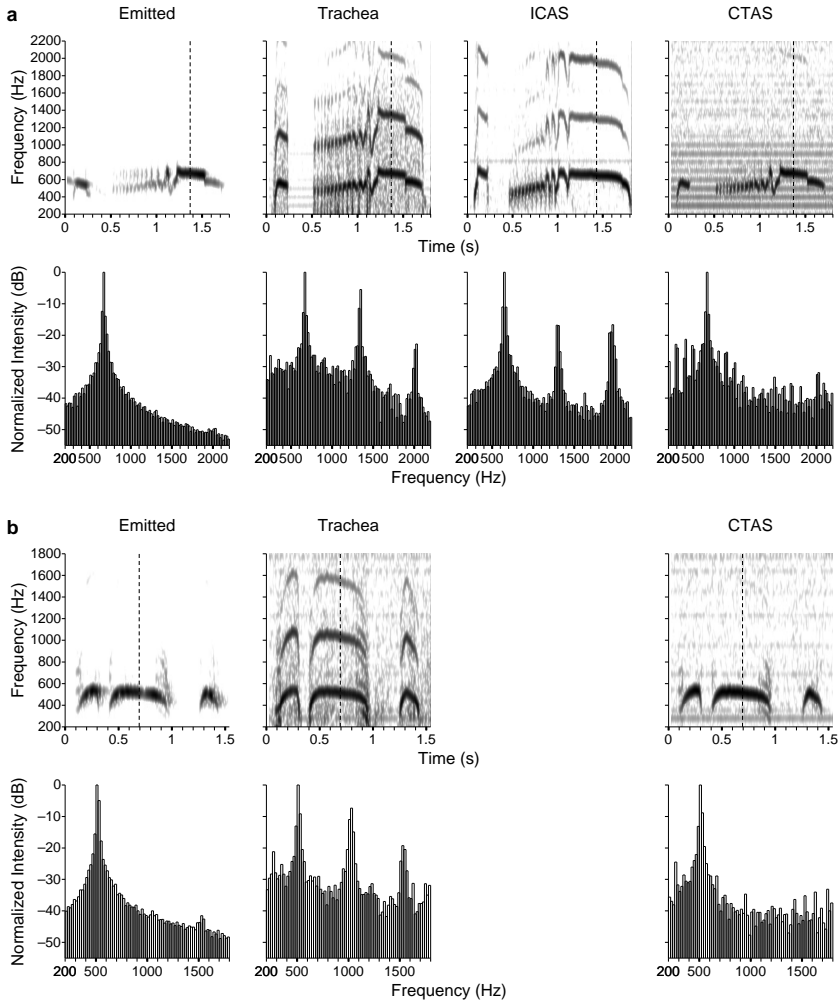


Figure 4.1: Sound signals in different compartments of respiratory and vocal system in a ring dove (a) and Eurasian collared dove (b). Emitted, tracheal and cranial thoracic air sac (CTAS) signals are from the same coo vocalization. The interclavicular air sac (ICAS) signal is from a different coo, but from the same individual. Top panels in each subfigure are spectrograms, and the bottom panels are power spectra (in averaged 20 Hz bins) of a 100 ms segment, centered around the time indicated with a dashed line in the spectrogram above. The intensity values in power spectra are relative to the intensity of f_0 . The smaller peaks at integer multiples of 100 Hz in the CTAS signal of the ring dove (a) are an artifact due to electronic noise.

Dove	Sound signal				N
	Emitted	Trachea	ICAS	CTAS	
CD1	+	+			392
CD2	+	+		+	2
RD1	+	+			86
RD2	+	+		+	9
RD2	+		+	+	208
RD3	+		+		88
RD4	+		+	+	39

Table 4.1: Recorded combinations of sound signals. Plus symbols in the same row indicate that signals have been recorded concurrently. Eurasian collared doves are designated as ‘CD’, and ring doves as ‘RD’, while different numbers refer to different individuals.

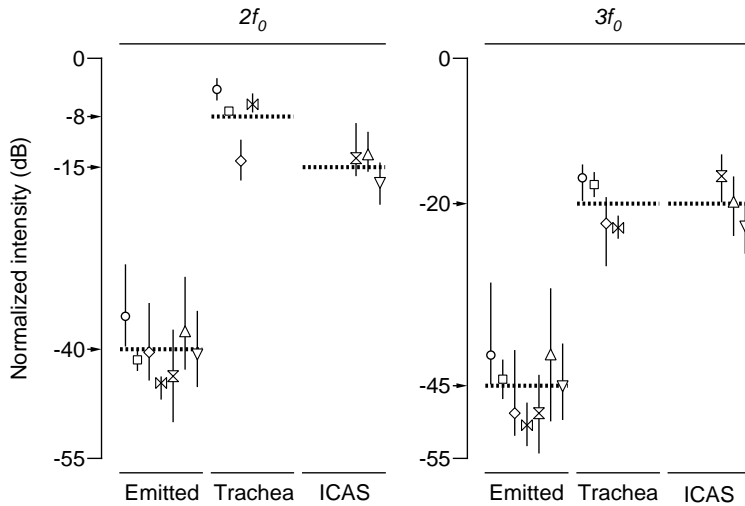


Figure 4.2: Intensities of second ($2f_0$) and third ($3f_0$) harmonics in emitted, tracheal and interclavicular air sac (ICAS) signals, relative to that of the fundamental (f_0). Symbols indicate means (with ranges specified by vertical lines), calculated from 20 recordings, except for the two cases in which less than 20 recordings were available (see Table 4.1). Identical symbols indicate that measurements are from concurrent recordings. Horizontal dashed lines indicate means per anatomical compartment, calculated from the mean values of individual birds. The values of these means are indicated through arrows on the vertical axis.

4.4 Discussion

Our results demonstrate that the source sound of pure-tone song in ring doves and Eurasian collared doves has a multi-frequency harmonic spectrum. The coo vocalization that is emitted by the doves consists only of the fundamental frequency part of this spectrum. We conclude therefore that the pure-tone quality of turtle-dove song originates by a source-filtering mechanism. Although we did not record source vibrations directly, but rather sound near the syrinx, the harmonic overtones in tracheal and air sac signals must be generated by the sound source, whatever its nature. Resonant structures can play an important role in sound production, but in sustained sounds they can only respond to frequencies at which they are excited, i.e. they cannot generate their own frequencies (Kinsler et al., 2000).

Both in the trachea and ICAS the spectral pattern of the first three harmonics is similar: the f_0 has the highest intensity, while the $2f_0$ and $3f_0$ show considerable, but decreasing, amounts of energy. It is interesting to note that this spectral pattern resembles that of human glottal flow waveforms during normal speech. Modal human speech source spectra decay according to f_0^{-2} (Flanagan, 1972). For $2f_0$ and $3f_0$, this gives intensities of -12 and -19 dB, respectively. In our doves these are, on average, -11 and -20 dB. It should be noted, however, that the accuracy of our intensity measurements is affected by three shortcomings. First, the thermistors used to record tracheal signals attenuate higher frequencies. This leads to a systematic underestimation of the intensity of $2f_0$ and $3f_0$, relative to f_0 . Because of the relatively low frequencies involved, this error is likely to be small. Second, thermistors, and to a lesser degree also piezoresistive pressure transducers, introduce energy at harmonic overtone frequencies into recorded waveforms due to their non-linear responses. This leads to an overestimation of the intensity of $2f_0$ and $3f_0$. Calibrations showed the levels of signal energy introduced this way to be far below the levels obtained in our experimental recordings. However, these calibrations were necessarily performed *ex situ*, in conditions different from those inside a vocalizing dove, and hence the real introduced error may deviate from that found in the calibration. Third, our measurements are based on signals recorded in the trachea and ICAS. This is close to the syringeal source, but not at the source. Hence the measured signals are inevitably affected by filtering, as illustrated by differences between the tracheal and ICAS signals (Fig. 4.2). Despite this uncertainty regarding the exact level of harmonics at the syringeal source, we believe the presence of high-intensity harmonic overtones close to the syrinx, measured by two very different types of measuring device, provides convincing evidence for the multi-frequency, harmonic nature of

this sound source.

We cannot directly extend our findings to pure-tone birdsong production in general, since we have only tested two closely related non-oscine species. It is possible that other species use different types of sound sources to generate pure-tone sound. Nevertheless, the results do show that it is not necessary to invoke source mechanisms radically different from valves, vibrating largely independently from resonant cavities, to explain pure-tonality. The idea of a source-filter production mechanism as originally proposed by Nowicki (1987) can adequately account for pure-tone birdsong.

Ballintijn and ten Cate (1998) concluded on the basis of helium studies that the pure-tone coo vocalizations of Eurasian collared doves are produced as such at the syringeal source. Our current study, however, shows this is not the case. Perhaps the interpretation of helium experiments depends too heavily on the assumption that vocal tract filtering can be modeled as the resonances of a simple tube. In reality, the avian vocal system includes a complex system of interconnected tubes and air sacs, partially separated by thin membranes. The resonant behavior of such a system is not easily predicted (Fletcher and Tarnopolsky, 1999).

Larsen and Goller (1999) recorded vibrations of lateral tympaniform membranes in the trachea concurrently with the emitted sound during brain stimulation-induced phonation in anaesthetized pigeons (*Columba livia*). Membrane vibrations had waveforms that were highly similar to that of the emitted sound, and were pure-tonal (f_0 exceeded $2f_0$ by at least 20 dB). This contrasts to the relatively strong tracheal harmonic overtones found in the pure-tone coos of our doves. It seems unlikely that the syringeal source spectrum in pigeons is fundamentally different from that in the dove species we studied. Since both doves and pigeons are in the family Columbidae and share a basically similar syringeal and vocal tract anatomy, we believe the discrepancy between our data and those of Larsen and Goller (1999) is likely due to the fact that the pigeon vocalizations were elicited from anaesthetized birds using brain stimulation.

The extent of filtering needed to transform the dove's multi-frequency source signal into a pure tone is remarkable. The relative intensity of $2f_0$ measured in the trachea is on average 32 dB higher than in the emitted coo. It has been suggested that the resonance characteristics of the trachea and oral cavity are the basis for vocal filtering (Gaunt and Nowicki, 1998; Nowicki, 1987; Nowicki and Marler, 1988; Hoese et al., 2000). If only tracheal resonance filtering were responsible for the generation of pure tones from a multi-frequency harmonic source, then we would expect to find just a pure tone inside the trachea during phonation. Our finding of harmonic overtones inside the trachea

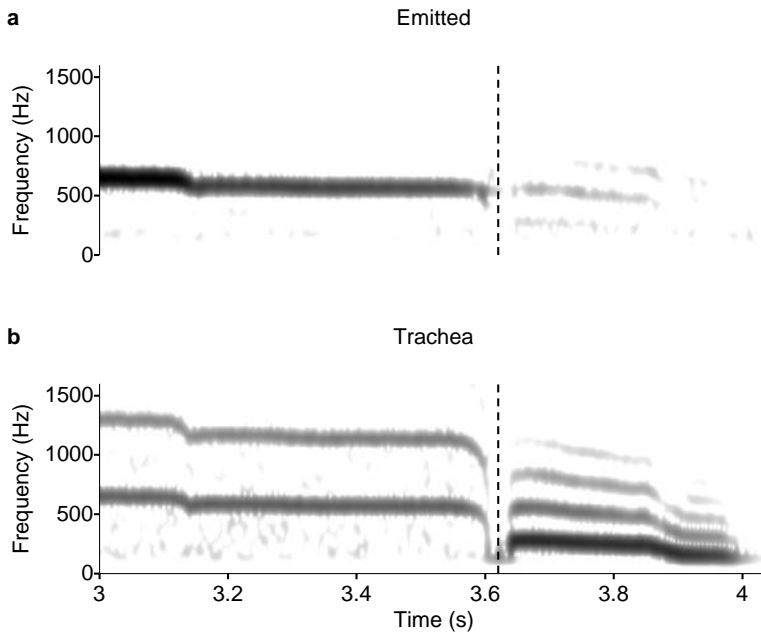


Figure 4.3: Spectrograms of emitted sound (a) and concurrent tracheal sound (b) at the end of a ring dove coo. The vertical dashed lines indicate where the coo (produced with the beak and nares shut) stops, and the inhalatory ‘wah’ begins.

shows therefore that in doves the trachea is not the only structure involved in filtering. This is perhaps not so surprising since a simple tube like a trachea resonates at a series of either odd and even, or only odd harmonic overtone frequencies, depending on whether or not the ends are closed (Kinsler et al., 2000). It is therefore not possible to have one of the harmonics of a multi-frequency source tuned to a sympathetic resonance frequency of a simple tube, and at the same time have all the other source harmonics filtered out by this tube.

Since ring doves (Gaunt et al., 1982), and most likely also collared doves, coo with their beaks and nares closed, one alternative explanation could be that the source sound is simply low-pass filtered below $2f_0$ when it passes through soft tissues as it radiates from the body. However, this hypothesis is inconsistent with the spectral structure of the ‘wah’ sound that is produced by both dove species during inhalation immediately following the end of a coo. This sound has multiple harmonics above a f_0 of 150–250 Hz. Although the intensity of f_0 is high in the trachea, it is almost completely absent in the

emitted vocalization (see Fig. 4.3). Simple low-pass filtering by transmission through soft tissue cannot account for this, since the frequency range is well below that of the coo. We therefore suggest that the filtering mechanism that makes coos pure-tonal involves a single-peak resonator that is sympathetic to the f_0 of the multi-frequency harmonic source. This would also explain the finding that coo intensity is strongly reduced when doves vocalize in helium (Ballintijn and ten Cate, 1998). A potential mechanism for this could be that of a Helmholtz resonator, in which compression of air in a cavity provides stiffness, while one or more connected air-filled tubes provide inertance and resistance (Kinsler et al., 2000). Cicadas employ a Helmholtz mechanism to produce loud, pure-tone sounds using a multi-frequency source (Bennet-Clark, 1999). In birds, the air sac system may be involved as the cavity part of such a resonator. This would provide an explanation for the existence of the medial tympaniform membranes, the function of which remains unclear since Goller and Larsen (1997a,b) showed they are not involved as primary sound generators. We suggest that these extremely thin membranes, situated in the wall of each bronchus, may acoustically couple the bronchi and trachea to the air sac system. Alternatively, or perhaps additionally, the air cavity of the inflated crop in cooing doves may serve as part of an acoustic resonator. But whatever the nature of the filter mechanism involved in pure tone bird-song, this study demonstrates its presence and raises the intriguing question why birds produce pure-tone song at all.

Acknowledgements We are grateful to Johan Bolhuis and Coen Elemans for comments on the manuscript and to S. Ronan for assistance with the experiments. This study was supported by a travel grant from the Netherlands Organization for Scientific Research (NWO) to GJLB, and by NIH grant NS29467 to RAS.

5 Mechanisms of frequency and amplitude modulation in ring dove song

Summary Birdsong assumes its complex and specific forms by the modulation of phonation in frequency and time domains. The organization of control mechanisms and intrinsic properties causing such modulation have been studied in songbirds, but much less so in nonsongbirds, the songs of which are often regarded as relatively simple. We examined mechanisms of frequency and amplitude modulation of phonation in ring doves, *Streptopelia risoria*. Spontaneous coo vocalizations were recorded together with concurrent pressure patterns in two different air sacs and air flow rate in the trachea. The results show that amplitude modulation is most likely the result of the cyclic opening and closure of a valve, instead of fluctuations in driving pressure as is the current explanation. Frequency modulation appears to be more complex than previously recognized, and consists of gradual, continuous time-frequency patterns, punctuated by instantaneous frequency jumps. Gradual frequency modulation patterns correspond to pressure variation in the interclavicular air sac, which is most likely the cause of such modulation. The cause of frequency jumps has not been identified, but can be explained on the basis of intrinsic properties of the vocal organ. Air sac pressure variation as a mechanism for frequency modulation contrasts with the specialized syringeal musculature of songbirds, and may explain why the fundamental frequency in nonsongbird vocalizations is generally modulated within a limited frequency range.

This chapter is also submitted for publication elsewhere as: Beckers, G. J. L., Suthers, R. A. & ten Cate, C. Mechanisms of frequency and amplitude modulation in ring dove song.

5.1 Introduction

The avian vocal organ, the syrinx, produces sound through vibrating labia or membranes, driven by respiratory air flow (Goller and Larsen, 1997a,b; Larsen and Goller, 1999). Birdsong assumes its complex and specific forms by the gating (on–off switching) of this sound and by the modulation of its properties both in time and frequency domains. This is achieved by coordinated neuromuscular activity of respiratory, syringeal and craniomandibular systems (for a review see Suthers et al., 1999), but part of the complexity may also emerge due to intrinsic, nonlinear characteristics of the syrinx (Fee et al., 1998; Fletcher, 2000). Elucidating the organization of such control mechanisms and intrinsic properties is a major challenge in the understanding of birdsong complexity.

Songbirds as a group are widely considered to have a higher level of vocal virtuosity and complexity than nonsongbirds. This difference is attributed to their more complex syringeal musculature (Gaunt, 1983). The involvement of different peripheral systems in the modulation of phonation has been relatively well-studied experimentally in a number of different songbird species (e.g. Ruppell, 1933; Suthers et al., 1994; Goller and Suthers, 1996; Fee et al., 1998; Hoese et al., 2000; Larsen and Goller, 2002). Studies on modulation in nonsongbirds, however, are almost nonexistent, although the gating of phonation and sound generation as such have received considerable attention (e.g. Youngren et al., 1974; Lockner and Murrish, 1975; Nottebohm, 1976; Gaunt et al., 1977; Brackenbury, 1980; Suthers and Hector, 1985). A better understanding of the differences in constraints and possibilities in phonatory control mechanisms of songbirds and nonsongbirds could provide insight in the factors that underlie the disparity in song complexity between these two groups. In addition, phonatory modulation mechanisms in nonsongbirds are interesting in their own right, since about half of all bird species are nonsongbirds, and vocal communication in these taxa seems to be no less important than in songbirds.

In the current study we address modulation of phonation in a nonsongbird, the ring dove, *Streptopelia risoria*. This species is particularly interesting because its vocalization behaviour has been studied before in the context of such diverse aspects of animal behaviour as reproduction (e.g. Lehrman, 1965), development (Nottebohm and Nottebohm, 1971), genetics (Lade and Thorpe, 1964), neuroendocrinology (e.g. Cheng et al., 1998), behavioural ecology (e.g. de Kort and ten Cate, 2001), and perception (Beckers et al., 2002a). Moreover, studies by Gaunt et al. (1982) on the mechanism of phonation and sound modulation in ring doves provide a basis for further study.

Ring doves are the domesticated form of the African collared-dove, *Streptopelia roseogrisea*. The vocalizations of domestic and wild forms are not different (Goodwin, 1983; Slabbekoorn et al., 1999), and they are considered to be the same species (Baptista et al., 1997). Ring dove coos have been described as relatively simple and stereotypic vocalizations (Nottebohm and Nottebohm, 1971; Slabbekoorn et al., 1999), and do not vary appreciably between birds (Goodwin, 1983). The development of normal coo vocalizations has been shown not to depend on learning (Nottebohm and Nottebohm, 1971), and their acoustic structure has been studied in more detail by Gaunt et al. (1982) and Slabbekoorn et al. (1999). Ring dove coos consist of two sound elements, separated by a silent pause (hereafter referred to as e_1 , e_2 and p), while their spectral structure consists only of a fundamental frequency (f_0) without any overtones. The first part of e_2 is amplitude-modulated (and is therefore seen as a separate note by Miller and Miller (1958)), which gives rise to a trill-type, rolling quality. Gaunt et al. (1982) and Miller and Miller (1958) report that there is little frequency modulation, although frequency varies slightly in the last part of e_2 . Slabbekoorn et al. (1999) did not give any specifics on frequency modulation patterns, but report that the frequency varies on average from 388 to 822 Hz within coos.

To identify physiological correlates of phonation and modulation, Gaunt et al. (1982) recorded ring dove coo vocalizations together with concurrent air pressure variation in the trachea and posterior thoracic air sac (PTAS; = caudal thoracic air sac), and electromyograms (EMGs) of syringeal and abdominal muscles. From this they concluded the following: (1) Syringeal muscles act to set the syrinx in vocalizing position, but they are probably not important for modulation. (2) The overall two-element coo phonation pattern is due to two large peaks of driving air sac pressure, caused by activity of abdominal muscles. (3) The trill-type amplitude modulation is due to pulsatile activity of the abdominal muscles, which cause an oscillation in driving air sac pressure. In addition, Gaunt et al. (1982) recognized a second, subtle type of amplitude modulation, which they explained by a muffler action of lateral tympaniform membranes (LTMS) and small differences in the vibration frequencies of median tympaniform membranes (MTMS). However, this explanation has become unlikely since Goller and Larsen (1997a) showed that not the MTMS but the LTMS are the sound source in domesticated rock pigeons, *Columba livia*, a species that is closely related to ring doves (Johnson et al., 2001), and has a very similar syringeal anatomy.

One goal of our current study was to get a more complete overview of physiological events during the trill-type amplitude modulation, which is a common phenomenon in dove vocalizations. Another goal was to examine

frequency modulation and its mechanistic basis. Despite the claim that frequency modulation in ring doves is limited (Miller and Miller, 1958; Gaunt et al., 1982), it does exist and is also present in many other dove species. Even small differences in frequency have been shown to have communicative meaning in the one species for which this has been tested, the Eurasian collared-dove, *S. decacoto* (Slabbekoorn and ten Cate, 1998; ten Cate et al., 2002). Insight in its mechanistic basis might also provide a better understanding of why frequency modulation is limited in the first place. To achieve this, we set out to supplement the measurements of Gaunt et al. (1982) by recording spontaneous coo vocalizations together with concurrent air pressures in the interclavicular air sac (ICAS) and cranial thoracic air sac (CTAS), and air flow rate in the trachea. The results lead us to conclude that the mechanism for amplitude modulation is different from the current model. We also propose a novel mechanism for frequency modulation.

5.2 Materials and Methods

Subjects We used five adult male ring doves (*S. risoria*) as subjects. Three doves were obtained commercially in the USA, and two in the Netherlands. Data recorded from four of the birds have also been used for a study on pure-tone sound production. Results from this study have been reported elsewhere (Chapter 4; Beckers et al., 2002b).

Surgical procedures and recording of data The procedure to record air sac pressure and tracheal flow rate is described in detail in Suthers et al. (1994). Therefore we will only give a summary here.

After birds were anesthetized with isoflurane (Abbott Laboratories), a midline incision was made in the skin between the clavicles to expose the trachea as it entered the ICAS membrane. Tracheal air flow was measured with a microbead thermistor probe (Thermometrics, BB05JA202N) inserted into the tracheal lumen, just rostrally to the interclavicular membrane. Thermistor wires were routed subcutaneously to connectors on a backpack that the birds wore. The flow rate in the trachea was measured by a feedback circuit in which the current needed to maintain the heated thermistor at a constant temperature was nonlinearly proportional to the rate of air flow (Hector Engineering).

Air sac pressure was measured by a piezoresistive silicone diaphragm pressure transducer (Fujikura FPM-02PG) attached to an air sac cannula consisting of a flexible silastic tube (Dow Corning, ID = 1.02 mm, wall thickness = 0.57 mm). A cannula 18.5 cm long was inserted into the ICAS through a small hole

in the interclavicular membrane. From there, it was routed subcutaneously to the backpack carrying the pressure transducer. The CTAS was cannulated by a similar tube 14 cm long inserted into the air sac through the abdominal wall just posterior to the last rib and a few mm lateral to the ventral midline. The cannulae extended 13 mm into the air sacs and tissue adhesive was used to insure an air-tight seal.

Vocalizations were recorded on a condenser microphone (Audio Technica AT835b or Sennheiser MKH 40) placed 0.5–1 m in front of the cage. All signals (emitted vocalization on microphone, ICAS pressure or tracheal flow rate, and CTAS pressure) were recorded digitally (20 kilosamples/s) on a rotary storage recorder (Metrum Information Storage, model RSR 512) or on a DAT data recorder (TEAC, model RD135T). Coo vocalizations were recorded with either concurrent tracheal flow rate and CTAS pressure signals, or with ICAS and CTAS pressure signals. We transferred the recorded signals from tape to a microcomputer by resampling (20 kilosamples/s), using a Data Translation DT-2821G board and a TTE J87 anti-aliasing filter (high cut-off at 8 kHz, stop-band attenuation 60 dB per 1/3 octave).

The pressure transducers were calibrated and showed a linear response for the ranges that we encountered in our recordings. During recording, however, we could not determine absolute pressure levels because of arbitrary dc-offsets introduced by the recording setup. Normally it is possible to deduce the atmospheric pressure level using air flow reversals in concurrent tracheal flow recordings. Ring doves, however, keep their beak and nares tightly closed during cooing, while expiratory air is deflected into the crop (Gaunt et al., 1982). Before and after phonation, flow reversal in the trachea often does not coincide with an opened beak (personal observation), and hence does not necessarily indicate when air sac pressure is atmospheric. Since we were mainly interested in how phonatory characteristics are associated with changes in air sac pressure, we made no further attempts to determine absolute pressure levels, but used relative pressures instead. For the analyses of tracheal flow rate, we were also restricted to the use of relative values, since the relationship between flow rate and thermistor output is nonlinear, and changes over time due to gradual mucus deposition on the thermistor tip.

Data analysis Recordings were analyzed with the software program Praat (available from Paul Boersma and David Weenink, www.praat.org), version 4.0.5 for Linux. Air sac pressure and tracheal flow rate signals were low-pass filtered digitally at 100 Hz, using a built-in function of the Praat program (frequency domain filter, Hann-like shaped band, 100 Hz smoothing). Although all vocalizations were recorded with a microphone, we used oscillations asso-

ciated with near field sound in the tracheal flow and oscillations in air sac pressure for analyses. These oscillations were retrieved by band-pass filtering flow and pressure signals at 350–800 Hz (50 Hz smoothing). The reason for preferring such signals is that they do not suffer from the sometimes severe acoustic artifacts that are introduced in microphone signals, if the recording room is not specifically designed for low-frequency sounds like dove coos (personal observation, see discussion).

We looked for correlates of gating and frequency modulation in tracheal flow rate and ICAS and CTAS pressure patterns by printing and visually comparing them together with spectrograms of concurrent coo vocalizations. This was done for all recordings. We did not differentiate between nest-, bow- and perch-coos since in ring doves there are no apparent differences in overall acoustic structure.

For quantification of relationships between flow and pressure patterns and the modulation of coo frequency, we focused on a selection of 10 coos for each combination of recorded variables. If possible, we selected vocalizations from different recording sessions, and further based the selection on a favourable signal-to-noise ratio. For one recording combination (tracheal flow with CTAS pressure in RD2), however, we only obtained eight coos recorded in a single session. For each coo we determined the fundamental frequency (f_0) in consecutive 3-ms time frames, using Praat's autocorrelation function (Boersma, 1993), and the mean values of available tracheal flow rate, and ICAS and CTAS signals in these frames. On average, this resulted in a series of 366 measurement sets per coo. All data were read into a matrix file, which was imported into SPSS for Windows, version 10.1, for statistical analyses.

Associations between the modulation of coo frequency and the three recorded physiological variables were examined quantitatively by computing product–moment correlation coefficients between concurrent 60-ms time segments (which thus consisted of 20 consecutive 3-ms frames). This was only done for time segments in which the f_0 time–frequency contour was continuous, so we excluded segments with silent intervals, and segments that contained frequency jumps (see Results). Very rarely, f_0 was almost constant within a time segment. Since correlation coefficients are not informative in such cases, we excluded segments in which the standard deviation of frequency was less than 1 Hz. Significance tests of the obtained correlation coefficients are not appropriate because samples are not statistically independent in time series. Also the strength of correlation should not be given much explanatory power, because there is no reason to assume that associations, if any, would be linear. In the case of flow rate, the measurements are known to be nonlinear in themselves, so a comparison of flow–frequency coefficient

Dove	Recorded signals				N
	Vocalization	Tracheal flow rate	ICAS pressure	CTAS pressure	
RD1	+	+		+	86
RD2	+	+		+	8
RD2	+		+	+	208
RD3	+		+	+	88
RD4	+		+	+	39
RD5	+	+		+	27

Table 5.1: Recorded combinations of sound signals. Plus symbols in the same row indicate that signals have been recorded concurrently. N denotes the number of recorded coos.

magnitudes with those of pressure variables makes a priori no sense. Instead, we just used the sign of correlation coefficients (positive or negative) as an indication of whether frequency and a particular physiological variable varied into the same direction within a 60-ms time segment or not. Per individual bird we categorized the correlation coefficients of all time segments into two categories: positive and negative coefficients. If continuous frequency modulation is consistently associated with one or more of the recorded physiological variables, then we would expect almost all of their coefficients to fall in only one of these two categories.

5.3 Results

Table 5.1 gives an overview of which signal combinations were obtained for each bird, as well as the number of recorded coos. Figures 5.1 and 5.2 give representative examples of the two types of signal combinations that we recorded concurrently. Visual comparison of the printed recordings showed that there is little variation in the overall patterns of flow and pressure within and between individuals. We will therefore address amplitude and frequency modulation of phonation using these figures. Qualitative descriptions are valid for all recorded signals, unless we state otherwise.

Amplitude modulation and gating Phonation is always accompanied by air flow through the trachea, while silent intervals within coos are always accompanied with a stop, or at least strong reduction, of air flow (Fig. 5.1a and b). This is also true for the amplitude-modulated part of e_2 , which appears to be

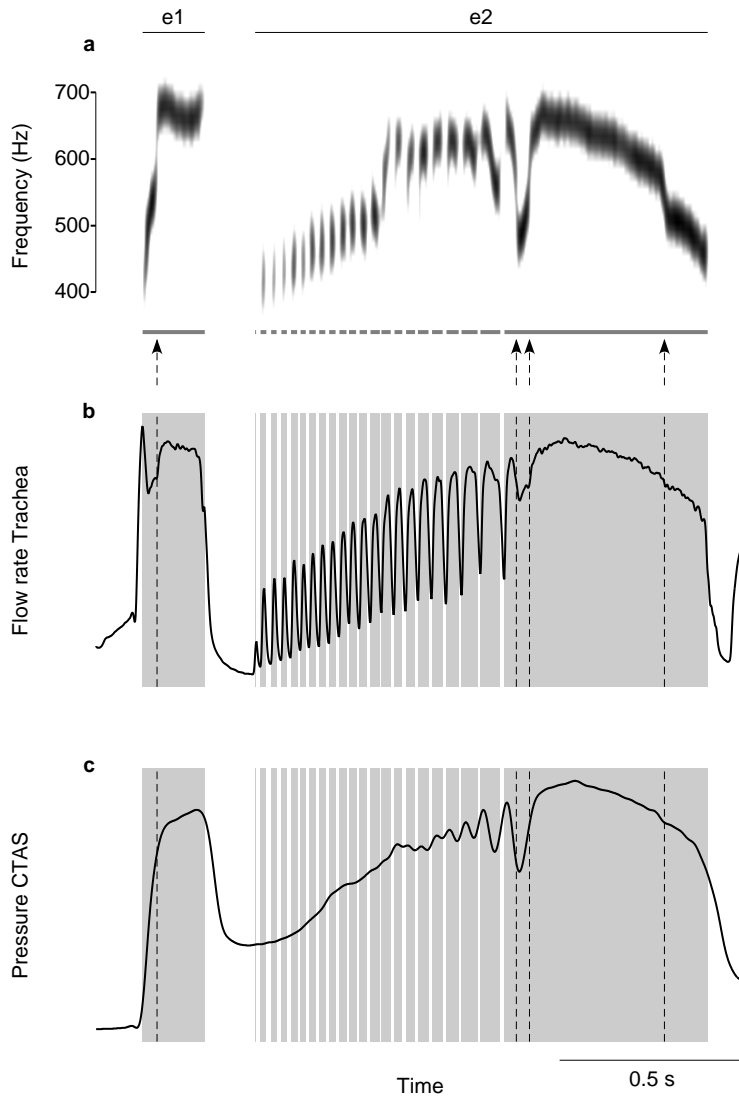


Figure 5.1: Concurrent vocalization (a), tracheal flow rate (b) and CTAS pressure (c) signals for one coo. Grey bars beneath the spectrogram and grey areas in the flow and pressure plots indicate where sound is produced, as measured from the sound oscillogram (not shown). Arrows and dashed lines indicate the location of frequency jumps. Recorded flow rate levels during the silent intervals of the amplitude-modulated part of e_2 may not reach zero because the microbead thermistor fails to track the very fast and large changes in flow rate.

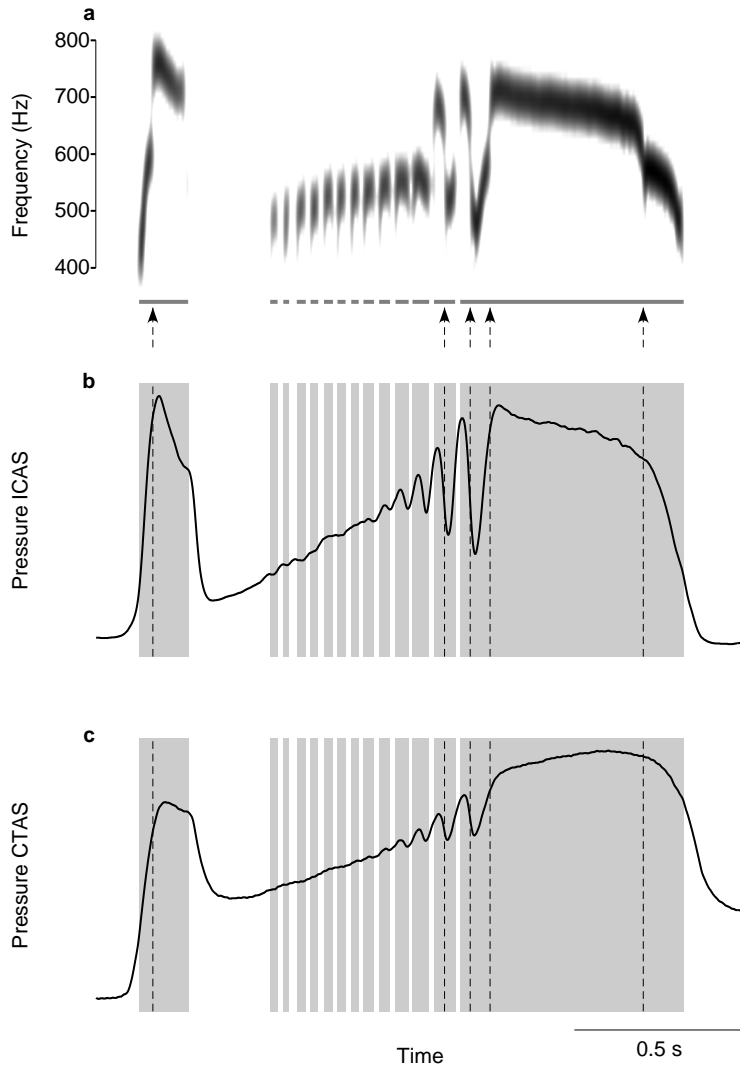


Figure 5.2: Concurrent vocalization (a), ICAS pressure (b) and CTAS pressure (c) signals for one coo. For details see Fig. 5.1.

a series of small sound elements separated by silent intervals (Fig. 5.3). Thermistors record air flow rate, but not the direction of air flow. Gaunt et al. (1982) have shown, however, that tracheal pressures are always much lower than air sac pressures during ring dove vocalizations, so all air flow during phonation (Fig. 5.1b) must be in an expiratory direction through the syrinx.

Pressure patterns in the CTAS (Figs. 5.1c and 5.2c) ICAS (Fig. 5.2b) look similar to the PTAS pressure pattern reported by Gaunt et al. (1982): pressures are high during phonation (e_1 and e_2) and reduced during the silent interval p . This overall pattern of high pressure during phonation and low pressure during silent intervals, however, does not hold for the amplitude-modulated part of e_2 . Here, both ICAS and CTAS pressures rise gradually during the first part, also when phonation and tracheal air flow stop (Fig. 5.3). In the last part of the amplitude modulation, pressures start to oscillate with increasing amplitude. Sometimes a slight oscillation of pressure can also be seen in the first part of the amplitude modulation, but always in a gradually rising pressure pattern. Pressure oscillations are always much stronger in the ICAS than in the CTAS, and continue for one cycle into the non-modulated part of e_2 . The pressure reduction phase of a cycle starts at or slightly after the start of a short sound element, and thus the start of a burst of air flow through the trachea.

Frequency modulation For the five doves tested, f_0 is modulated within an on average 354 Hz wide frequency band. The mean centre of this band is at 563 Hz, so f_0 is modulated for about one octave in ring dove coos. Frequency modulation patterns appear to be considerably more complex than previously recognized.

All coos show two types of frequency modulation: continuous frequency modulation and frequency jumps. Continuous frequency modulation is characterized by a gradual change of f_0 over time. At frequency jumps, the gradual f_0 time–frequency contour is momentarily disrupted, as f_0 almost instantaneously jumps to a different frequency range (Fig. 5.4). Phonation, however, is not interrupted at frequency jumps, although changes in amplitude can often be observed. The arrows and dashed lines in Figs. 5.1a and 5.2a indicate examples of frequency jumps in complete coo vocalizations. The occurrence and timing of jumps, however, can vary considerably between coos.

Frequency jumps often go together with a sudden, small increase or decrease in tracheal flow rate, but the directions of frequency and flow rate change are not always the same. We did not observe any systematic changes in air sac pressures that can be linked to frequency jumps.

Continuous frequency modulation appears to correspond very well to patterns of ICAS pressure change (for example compare Fig. 5.2a and b). The

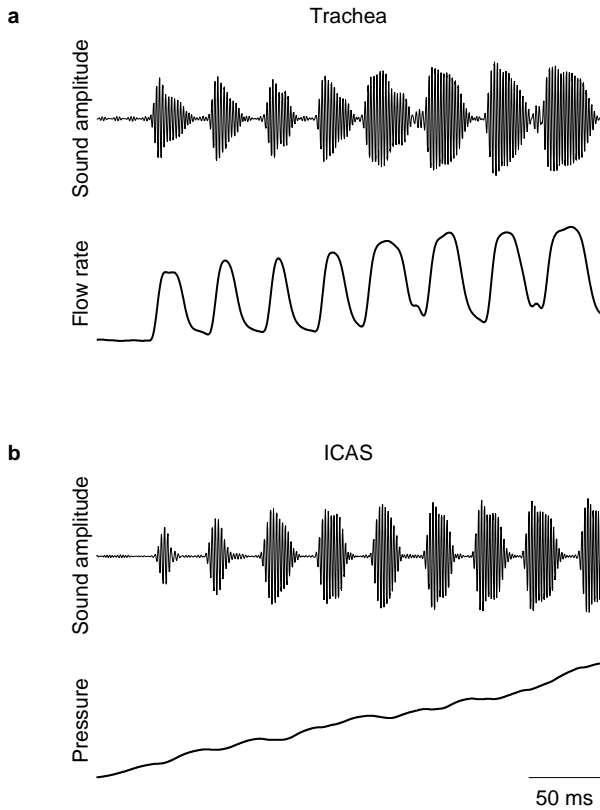


Figure 5.3: Gating of sound in the first part of the amplitude-modulation in *e2*, recorded in trachea (a) and ICAS (b). Recordings (a) and (b) are from different coos, but originate from the same individual. Both sound oscillograms in the trachea (a) and ICAS (b) show that phonation is completely interrupted by silent intervals of 10 ms or more. This modulation pattern corresponds to the pattern of flow rate in the trachea, which is reduced to zero or near-zero during silent intervals. ICAS pressure gradually increases in this part of the amplitude modulation, although the rate of increase is sometimes reduced during the sound intervals. CTAS pressure patterns are similar to those in the ICAS in this part of the coo, and are therefore not shown.

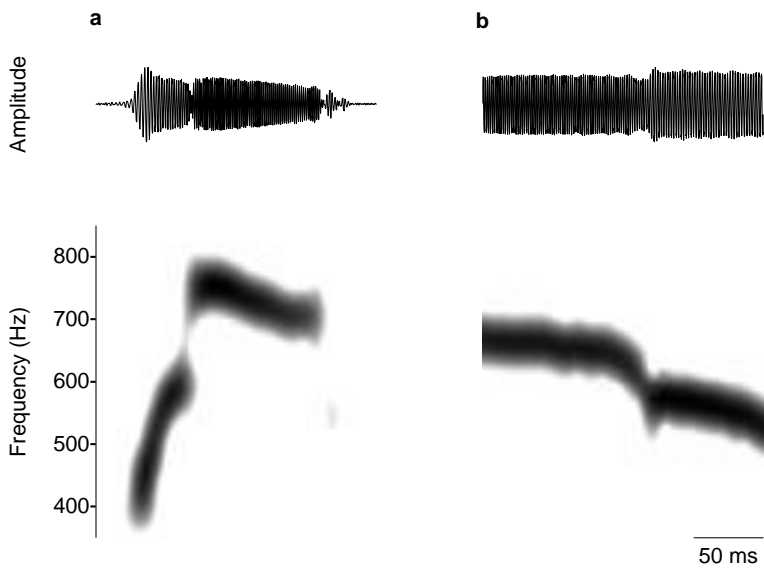


Figure 5.4: Two examples of a frequency jump. The frequency gap crossed by a jump varies within and between coos. A relatively large jump is shown in (a), in which f_0 jumps from about 600 to 750 Hz within 10 ms, and a smaller one in (b), in which f_0 jumps from 600 to 550 Hz in a similar time span. During these phenomena, the generation of sound is not interrupted, as seen in the oscillograms. Spectrogram settings: time step 2 ms, frame length 20 ms, dynamic range 15 dB.

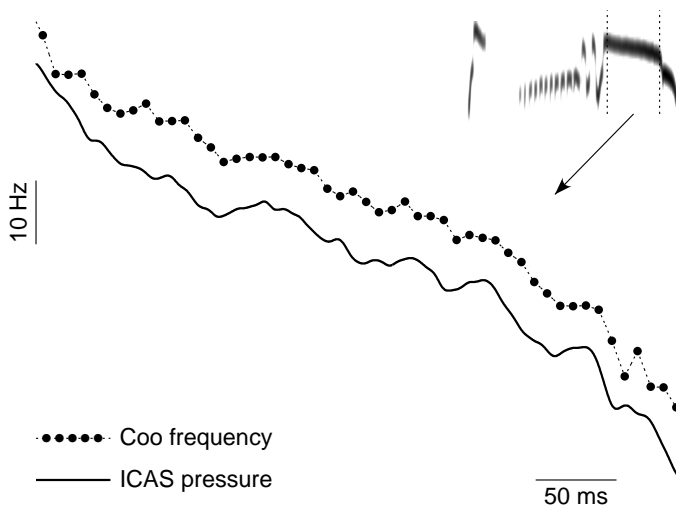


Figure 5.5: Detail of concurrent frequency modulation and ICAS pressure patterns in a segment of *e2*. The position of the segment within the coo is indicated in the spectrogram (top right corner). Frequency of phonation was determined at 8 ms intervals with an autocorrelation algorithm (Boersma, 1993) using the microphone-recorded vocalization. Signals have not been adjusted horizontally for a very small time delay difference.

visual examination of all obtained recordings revealed that sound segments not interrupted by silent intervals or frequency jumps can always be scaled to match concurrent ICAS patterns. In low-noise recordings, the close correspondence between ICAS pressure and f_0 modulation can often be seen to hold on a detailed scale (Fig. 5.5).

To verify whether or not ICAS pressure is consistently positively associated with f_0 modulation, we created scatterplots of ICAS pressure and coo frequency for each individual coo (three doves, 30 coos, e.g. Fig. 5.6a–c). All these scatterplots showed the same pattern: there is a positive, and predominantly linear relationship between ICAS pressure and coo frequency within continuous f_0 segments. At frequency jumps, the slope of this relationship usually changes somewhat. The close correspondence between ICAS pressure and frequency modulation only breaks at the onset and offset of phonation in the short sound pulses of the amplitude modulation. However, these tran-

sients are maximally 10 ms in duration, and constitute less than 5% of the coo sound. Moreover, scatterplots show that when this occurs, the overall relationship in amplitude-modulated parts remains linear and positive (see for example Fig. 5.6a, in which the yellow and light-green parts are from the amplitude-modulation).

Tracheal flow rate and CTAS pressure often also positively correspond to continuous frequency modulation patterns, but not always. Sometimes they even go opposite ways for the larger parts of $e1$ or $e2$ (compare for example the frequency pattern with CTAS signals in Figs. 5.1 and 5.2). An examination of scatterplots of the focal coo selections showed indeed that most tracheal flow rate and CTAS pressure plots contain smaller or larger segments that are of opposite directions (three doves, 28 coos, and 5 doves, 58 coos, respectively, e.g. Fig. 5.6d–f and 5.6g–i). Frequently, a switch from positive to negative relationships, or vice versa, can even be seen to occur within continuous segments of phonation.

Figure 5.7 shows that a quantitative analysis of the focal coo selections reinforces our qualitative observations. In 99.8% of all 60-ms segments of continuous phonation, correlation coefficients of ICAS pressures with f_0 are positive ($N = 332$, three doves). For tracheal air flow and CTAS pressure this is only 70.1% and 67.1%, respectively ($N = 221$, three doves, and $N = 553$, five doves), and hence the association between f_0 and these variables is not consistent.

5.4 Discussion

The main conclusions from our study are as follows: (1) The overall pattern of f_0 modulation in ring dove coos is strongly and positively associated with the change of air pressure in the ICAS. (2) The overall continuous and gradual f_0 contour is punctuated by instantaneous frequency jumps, not related to any observable changes in pressure or flow patterns. (3) The trill-type amplitude modulation present in ring dove coos is not caused by an oscillation of sub-syringeal pressure, as previously believed (Gaunt et al., 1982), but rather by the repeated opening and closure of an as yet unidentified valve.

The strong and positive association between f_0 modulation and ICAS pressure is likely to be a causal one. It has been shown in a related (Johnson et al., 2001) columbid species, *Columba livia*, that the source of coo vocalizations are the vibrating lateral tympaniform membranes (LTMS), located in the lateral walls of the tracheal part of the syrinx (Goller and Larsen, 1997b; Larsen and Goller, 1999). Phonation frequency is ultimately determined by these membranes' resonance frequencies, which depend on their geometric charac-

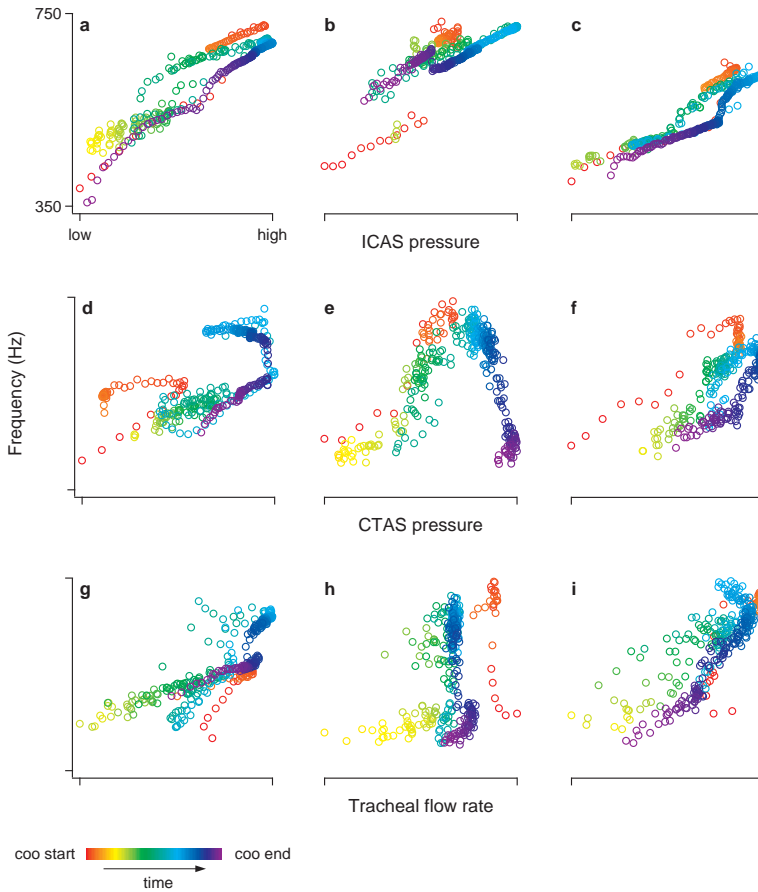


Figure 5.6: Coo frequency plotted against ICAS pressure, CTAS pressure and tracheal air flow. Each plot represents a single, complete coo, and circles indicate value pairs as measured in 3-ms time frames of that coo. Circles are coded on a colour scale so that the development over time within a coo can be traced. Each row contains three representative coos, originating from different individuals.

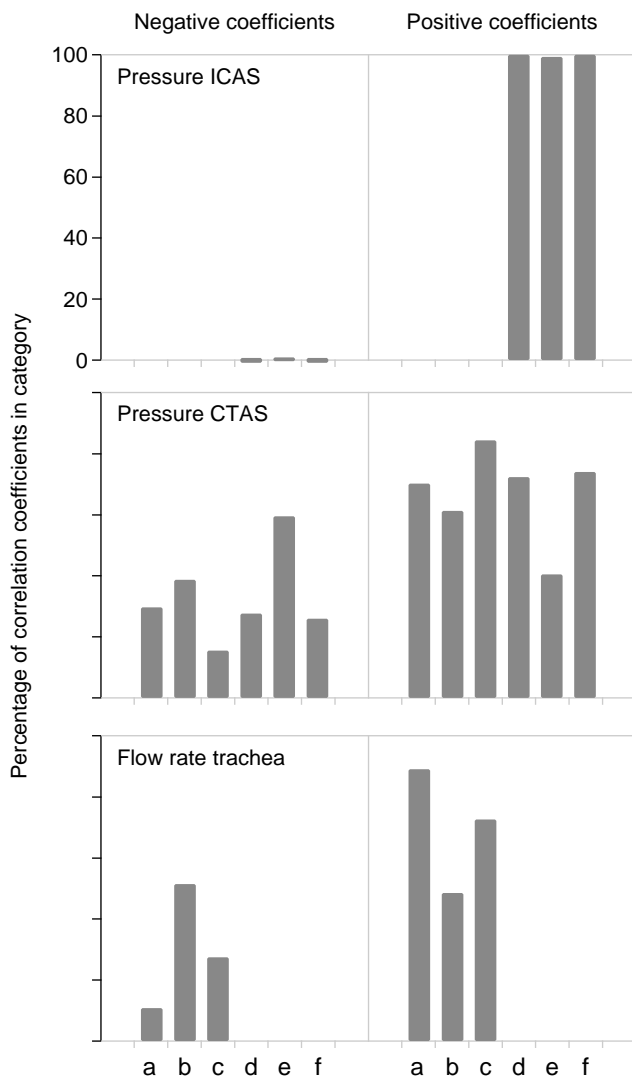


Figure 5.7: Association of coo fundamental frequency with tracheal flow rate, and ICAS and CTAS pressure. a = RD5; b = RD1; c = RD2; d = RD2; e = RD3; f = RD4 (see Table 5.1).

teristics, material density, and the tension applied to them (Fletcher, 1992). The syrinx is located within the ICAS, and the LTMs are in direct contact with ICAS space (King, 1989). It is therefore plausible that variation in ICAS pressure directly modulates LTM tension, and thus the frequency of phonation.

Hypothetically, it remains possible that it is not ICAS pressure that modulates f_0 , but a third factor that causes both the change of frequency and ICAS pressure. The only likely candidate to be included in such a relationship would be the tracheolateral muscles, the caudal end of which directly insert on the LTMs. These muscles could vary the degree of LTM adduction, and therefore at the same time both LTM tension and syringeal resistance. A change in syringeal resistance could in turn result in an increase of tracheal flow, and a reduction of subsyringeal (and consequently ICAS) pressure. If this were the case, however, we would also expect a consistent association of coo frequency and tracheal air flow rate, given our results that show a consistently positive association with ICAS pressure. But our results show that flow rate is not consistently associated with f_0 modulation, and hence we can reject this alternative hypothesis.

The occurrence of sudden frequency jumps in bird vocalizations has been reported earlier for Eurasian collared doves, *Streptopelia decaocto*, (Gürtler, 1973; ten Cate, 1992; Ballintijn and ten Cate, 1998), which are closely related to ring doves (Johnson et al., 2001). In this species, frequency jumps have been shown to have communicative meaning (Slabbekoorn and ten Cate, 1998; ten Cate et al., 2002). Frequency jumps have also been reported to occur in vocalizations of zebra finches, *Taeniopygia guttata*, where their mechanistic origin is attributed to mode locking in the syringeal dynamics (Fee et al., 1998). Mode locking happens when a nonlinear interaction constrains two oscillating components of a system to maintain a small integer ratio of frequencies. Mode-locking transitions may occur because the characteristic frequency of one component is changed relative to the other, and the oscillation frequency suddenly jumps to achieve a new stable integer ratio. We believe that mode-locking transitions are also a likely explanation for frequency jumps in ring doves, because other types of dynamical behaviour which the nonlinear interaction of two oscillatory components can lead to, i.e. sudden transitions to subharmonic and chaotic phonation (Wilden et al., 1998), are apparent in the normal vocalizations of related dove species (such as *S. tranquebarica* and *S. orientalis*), and occasionally in aberrant ring dove coos (personal observations).

The existing model for the trill-type amplitude modulation in ring dove coos, namely an oscillatory driving air sac pressure caused by the pulsatile action of abdominal muscles (Gaunt et al., 1982), is not compatible with our

results. We did find oscillations in air sac pressures, but only during the last part of the amplitude modulation. A close re-examination of the figures published by Gaunt et al. (1982) shows that this is also the case in PTAS pressure. Amplitude modulation is already present well before air sac pressures start to oscillate, when air sac pressures gradually increase, but tracheal air flow and phonation nevertheless stop in cyclic intervals (Fig. 5.3). We can only explain this pattern by the closure of some kind of valve. Amplitude modulation by the opening and closing of a valve also occurs in brown thrasher, *Toxostoma rufum*, song (Suthers et al., 1994). In ring doves, such action could for instance be performed by completely adducted LTM membranes, or by the glottis.

The oscillation of air sac pressures could be caused by active motor gestures of abdominal muscles, perhaps to achieve frequency modulation and add more complexity to the coo. However, a simpler explanation is that pressure oscillations are a consequence of the amplitude-modulating valve action. The ‘burst’ of air flow at the onset of a sound pulse, immediately after a silent interval with reduced or zero flow, may momentarily cause a reduction in subsyringeal air pressure, which, after a time lag, is compensated for by the action of abdominal muscles. Three observations fit this idea. (1) The reduction phase of a pressure oscillation takes place at, or first after, the start of a sound pulse. (2) Oscillations become significant only at sufficiently high flow rates and pulse durations, and the amplitude of the pressure oscillations increase as both air flow rate and duration of a sound pulse increase (Fig. 5.1). (3) Gaunt et al. (1982) report with respect to the pressure oscillation: “Abdominal muscular activity is highest during dips in the pressure curve. Each EMG pulse begins shortly after pressure begins to drop from a maximum and continues until pressure again reaches the level at which activity began.” Such EMG pattern suggests a compensatory action of expiratory muscles against the sudden pressure drop. It is possible that this compensation is not controlled by a central motor program. Suthers et al. (2002) recently showed that somatosensory feedback to expiratory muscles stabilizes subsyringeal pressure against fluctuations during vocalization in Northern cardinals, *Cardinalis cardinalis*. A similar mechanism might operate in doves.

Gaunt et al. (1982) recognized two types of amplitude-modulation in ring doves: the trill-type amplitude modulation that we addressed above, and a more subtle, beat-like amplitude modulation, which we did not address in this study. We too have observed a phenomenon of partial or complete beat cycles occurring, apparently unsystematically, in all parts of the coo, but only so in microphone recorded vocalizations. Phonation signals in tracheal flow or air sac pressure recordings never show this phenomenon. Hypothetically it is possible that such modulation originates from interaction of the sound

source signal with a resonance filter elsewhere in the animal. Resonance filtering of source signals has recently been demonstrated in ring doves (Chapter 4; Beckers et al., 2002b). However, interference through reflected sound waves and the formation of standing waves in enclosed rooms can also generate beat-like amplitude modulation (Kinsler et al., 2000). To test whether room acoustics could indeed introduce such strong modulations as observed by us and reported by Gaunt et al. (1982), we created synthetic ring dove coos with normal frequency modulation but constant amplitude, using the analytic signal technique as described by Mbu Nyamsi et al. (1994). Analysis of such signals generated by a speaker revealed that indeed strong beat-like amplitude modulations had been introduced, even when recorded in 'acoustic chambers' padded with standard acoustic foam. Recordings outside in the open field did not show such beat phenomena. Because we could not distinguish between beat-like amplitude modulation produced by the dove, if any, and that caused by the acoustics of the room, we decided not to investigate this phenomenon any further.

Our spectrograms of ring dove coos (Figs. 5.1 and 5.2) show that frequency modulation is much more complex than previously reported (Nottebohm and Nottebohm, 1971; Gaunt et al., 1982). Undoubtedly, a major reason for this difference in interpretation is the current availability of superior, digital techniques for spectrographic representation. An additional reason may be that frequency ranges in birdsong are often viewed spectrographically on an absolute and linear frequency scale, usually spanning a considerable part of our hearing range. Frequency modulation between 0.4 and 0.8 kHz does not appear impressive on such a scale. Perceptually, however, both to humans (e.g. Moore, 1997) and birds (e.g. Dooling, 1982), it is more appropriate to consider f_0 variation in terms of proportions. Our measurements show that from such a point of view the difference between the minimum and maximum f_0 in individual ring doves is about a factor of two (one octave), which is not negligible. Acoustic analyses of turtle-dove perch-coos by Slabbekoorn et al. (1999) show that this is also typical for the other 15 species in the *Streptopelia* genus. Moreover, even limited frequency modulation has important communicative function in the Eurasian collared dove, the sister species of the ring dove (Slabbekoorn and ten Cate, 1998; ten Cate et al., 2002).

Gardner et al. (2001) recently showed that much of the complexity of the song of canaries, *Serinus canaria*, can be modelled by smooth and simple variations of only two parameters. Despite their complexity, modulation patterns in ring dove coos may also arise from only two relatively simple, centrally coordinated motor variables. First, simple and smooth air sac pressure gestures determine the overall phonation and continuous frequency modulation

patterns. And second, amplitude modulation is caused by the cyclic gating action of a valve. Much of the remaining complexity, such as oscillatory frequency modulation and frequency jumps, can be explained on the basis of mechanisms that require no central control: feedback and intrinsic nonlinear properties of the syrinx.

What is the significance of our findings with respect to other bird species? We provide, to our knowledge for the first time, a mechanistic explanation for frequency modulation of phonation in a nonsongbird. ICAS pressure may also modulate frequency in the vocalizations of other nonsongbirds, given the fact that LTMS occur in many nonsongbirds (King, 1989), and assuming that the finding of LTMS as the sound generator in a pigeon and a parrot species (Goller and Larsen, 1997b) can be extended to other nonsongbirds. Frequency modulation in songbirds (Miskimen, 1951; Goller and Suthers, 1996; Larsen and Goller, 2002), and possibly parrots (Larsen and Goller, 2002), is achieved through the action of specialized syringeal musculature. A mechanism of frequency modulation by ICAS pressure variation could explain why many nonsongbird species often exhibit rather limited frequency modulation, on an absolute scale, as compared to songbirds. Muscles that modulate ICAS pressure can only indirectly vary tension of the sound generating structures and are not specialized in this task, as modulation of ICAS pressure is also essential for, and possibly constrained by, other physiological functions such as respiration. Moreover, the level of fluid power during bird vocalization depends on the level of air sac pressure (Brackenbury, 1977), and is positively correlated with sound intensity for the species investigated (Suthers and Goller, 1997; Gaunt et al., 1976). The overall modulation of frequency and sound intensity are thus likely to be coupled in ring doves, and possibly other nonsongbirds. A future study should investigate the possibility that maximizing vocal intensity, an important property of long-distance signals, may constrain frequency modulation.

Acknowledgements We are grateful to Selvino de Kort, Albertina Leitão, Hans Slabbekoorn, and Nienke Terpstra for discussion, and to S. Ronan and P. Snelderwaard for assistance with the experiments. This study was supported by a travel grant from the Netherlands Organization for Scientific Research (NWO) to GJLB, and by NIH grant NS29467 to RAS.

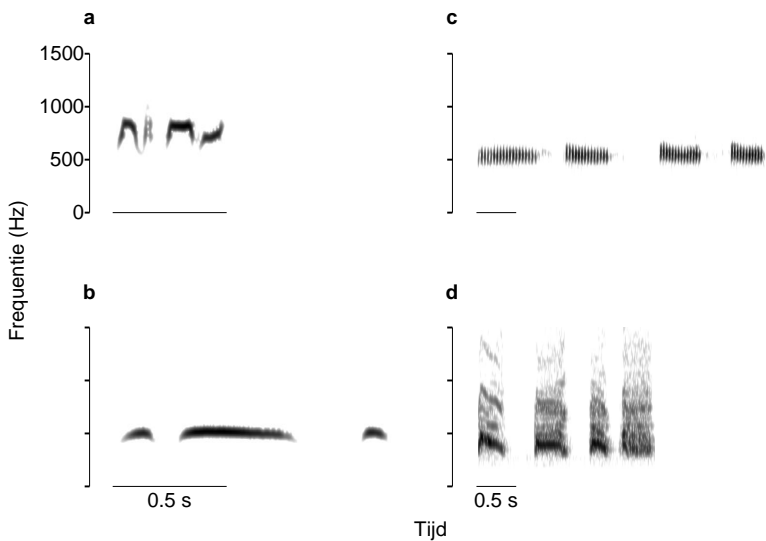
Nederlandse samenvatting

Vocale productie en perceptie in *Streptopelia* duiven

Van alle communicatieve geluiden die door dieren worden gemaakt zijn die van vogels waarschijnlijk het bekendst en meest opvallend. De belangrijkste redenen hiervoor zijn dat vogelgeluiden bijna overal te horen zijn, dat ze erg complex kunnen zijn, en dat er een buitengewoon grote diversiteit aan is. Wereldwijd komen er meer dan 9000 soorten vogels voor, die bijna allemaal via soortspecifieke geluiden communiceren. Meestal hebben individuele vogels ook nog verschillende vocalisaties, soms oplopend tot meer dan honderd bij één enkele vogel.

De alomtegenwoordigheid, complexiteit en diversiteit van vocale communicatie bij vogels vormen op zichzelf al een sterke motivering voor wetenschappelijk onderzoek naar dit type diergedrag. Waarom zijn er zoveel soorten vocalisaties? Hoe zijn ze ontstaan? Waar worden ze voor gebruikt? Waar komt hun complexiteit vandaan? Er is echter nog een andere reden voor onderzoek naar vogelzang. Voor de productie van vogelzang is snelle, complexe en gecoördineerde activiteit nodig van diverse neurale en perifere systemen. In dit opzicht lijkt vogelzang op menselijke spraak. Daarnaast is het zo dat in ongeveer de helft van alle soorten, vogels de patronen van hun vocalisaties moeten leren van soortgenoten. Ook in dit opzicht vertonen vogelzang en menselijke spraak parallellen. Vogelzang vormt een uniek modelsysteem voor onderzoek naar de manier waarop neurale en perifere systemen complexe vocalisaties genereren, in een context die lijkt op die van menselijke spraak.

In dit proefschrift onderzoek ik geluidsproductie- en perceptiemechanismen in *Streptopelia* duiven (tortelduiven). Elk van de 17 soorten *Streptopelia* duiven heeft een eigen, soortspecifiek koergeluid: de perch-koer. Het primaire doel van mijn onderzoek is beter te begrijpen hoe de eigenschappen van productie- en perceptiemechanismen de evolutie van deze communica-



Figuur 5.8: Spectrogrammen van perch-koeren van (a) *S. vinacea*, (b) *S. decaocto*, (c) *S. turtur*, and (d) *S. orientalis*. In een spectrogram staat de frequentie (=trillingssnelheid) van tooncomponenten uitgezet tegen de tijd. De intensiteit van deze tooncomponenten wordt weergegeven met een grijsschaal. Hoe donkerder een component is, des te hoger is de intensiteit.

tieve signalen kunnen beïnvloeden. Mijn werk maakt deel uit van een groter onderzoeksprogramma waarmee geprobeerd wordt te achterhalen hoe de koergeluiden evolutionair ontstaan zijn en welke factoren en processen ervoor gezorgd hebben dat ze hun huidige structuur gekregen hebben. Naast dit primaire doel dat in feite gericht is op het verklaren van het ontstaan van de specifieke diversiteit aan vocalisaties, hoop ik ook te kunnen bijdragen aan een beter begrip van hoe vogels in staat zijn de complexe patronen in hun vocalisaties te genereren.

De koeren van *Streptopelia* duiven verschillen akoestisch gezien in een verscheidenheid aan parameters. Deze verschillen kunnen in vier typen onderverdeeld worden.

Het eerste type verschil is de *temporele structuur*. Dit is het tijds patroon dat gevormd wordt door aan- en uitschakelen van geluid (dus het tijds patroon van geluidselementen en de tussenliggende pauzes). In figuur 5.8a-d zijn spectrogrammen van koergeluiden met verschillen in temporele structuur te zien. Twee van deze koeren hebben drie elementen (5.8a, b) en de andere twee hebben vier elementen (5.8c, d). Ook de duur van de elementen

en pauzes variëren. Zulke verschillen hebben tot gevolg dat elke koersoort zijn eigen ritme heeft.

Een tweede type verschil is de aanwezigheid van *amplitudemodulatie*. Amplitudemodulatie is de verandering van geluidsintensiteit in de tijd. Een deel van de *Streptopelia* soorten heeft een typische vorm van snelle amplitudemodulatie die leidt tot een trilachtig geluid. In figuur 5.8c is een voorbeeld te zien van een koer met zo'n type amplitudemodulatie, waarin elk van de vier elementen bestaat uit een snelle opeenvolging van geluidspulsjes. Bij de andere koeren (5.8a, b en d) is dit type amplitudemodulatie afwezig. Niet alleen het wel of niet voorkomen van amplitudemodulatie kan variëren tussen koersoorten, maar ook de precieze structuur ervan. Sommige koeren hebben bijvoorbeeld een snellere tril dan andere.

Een derde type verschil dat voorkomt tussen koeren is het patroon van *frequentiemodulatie*. De frequentie ofwel trillingssnelheid van een geluid bepaalt de toonhoogte. Frequentiemodulatie leidt dus tot een specifiek patroon van toonhoogteverandering in de tijd. Figuren 5.8a en 5.8b zijn voorbeelden van twee koergeluiden die duidelijk verschillen in het patroon van frequentiemodulatie. In tegenstelling tot de vocalisaties van veel zangvogels, zijn de koergeluiden van duiven en de meeste andere niet-zangvogels slechts in beperkte mate gemoduleerd in frequentie. Desondanks kan het patroon van modulatie wel communicatieve betekenissen hebben voor deze vogels.

Het laatste type parameter is de *tonale structuur*. De tonale structuur van een geluid wordt bepaald door het patroon van de tooncomponenten waaruit het is opgebouwd. Figuren 5.8a-c geven koeren weer met de meest eenvoudige tonale structuur: op elk punt in de tijd is er slechts sprake van één enkele tooncomponent. Zo'n geluid wordt ook wel een puur-tonaal geluid genoemd. De meeste *Streptopelia* koeren zijn puur-tonaal, maar sommige hebben een complexere tonale structuur. Het koergeluid van figuur 5.8d, bijvoorbeeld, heeft een complex patroon van vele tooncomponenten. Een verdere vergroting van de complexiteit in dit voorbeeld ontstaat doordat de tonale structuur ook nog eens verandert in de tijd (vergelijk bijvoorbeeld het eerste en het laatste element).

Dit proefschrift

In dit proefschrift beschrijf ik vier studies naar mechanismen van vocale communicatie bij *Streptopelia* duiven. Met mijn onderzoek probeer ik er achter te komen hoe belangrijk de akoestische verschillen tussen koersoorten zijn, in de perceptie van duiven zelf. Dit is interessant omdat het inzicht verschaft in welke koereigenschappen onderhevig kunnen zijn aan selectie door het ge-

drag van de signaalontvanger. Daarnaast onderzoek ik hoe de akoestische karakteristieken van *Streptopelia* vocalisaties fysiologisch tot stand komen. Kennis over de productiemechanismen die ten grondslag liggen aan koergeluiden leidt wellicht tot meer inzicht in de mogelijkheden voor, en waarschijnlijkheden van, evolutionaire veranderingen in deze vocale signalen.

In hoofdstuk 2 en 3 worden experimenten beschreven die inzicht verschaffen in welke aspecten van koergeluiden perceptueel belangrijk zijn voor duiven. In beide studies werden vogels met behulp van voedselbeloning getraind om te reageren op geluiden uit een bepaalde categorie, maar juist niet te reageren op geluiden uit een andere categorie. Vervolgens is bekeken hoe vogels reageerden op teststimuli die voor een deel lijken op trainstimuli uit de ene categorie, en voor een deel op die uit de andere, zodat kon worden afgeleid welke akoestische eigenschappen door de vogels gebruikt werden bij het maken van het onderscheid.

In hoofdstuk 2 kijk ik naar de perceptuele relevantie van soortspecifieke verschillen in temporele structuur en amplitudemodulatie. Met akoestische analyses en statistische technieken is eerder vast komen te staan dat de perchoeren van de verschillende *Streptopelia* soorten het best onderscheiden kunnen worden op grond van hun temporele structuur. Maar het feit dat de verschillen tussen soorten het grootst zijn in temporele structuur wil nog niet zeggen dat duiven deze ook daadwerkelijk gebruiken voor het maken van het onderscheid. Er zijn ook grote verschillen in amplitudemodulatie tussen soorten, en voor het menselijk oor zijn deze verschillen opvallender dan die in temporele structuur. In dit experiment dienden synthetische stimuli als model voor de natuurlijke koeren van twee *Streptopelia* soorten. Deze stimuli verschilden alleen in temporele structuur en amplitudemodulatie. Nadat duiven getraind waren om deze stimuli te onderscheiden, werd hun respons op hybride geluiden getest, waarbij telkens één parameter gevarieerd werd. Uit de resultaten blijkt dat de meeste vogels beide parameters gebruiken om de geluidsstimuli te classificeren, maar dat amplitudemodulatie iets belangrijker is.

Een groot voordeel van het experimentele design in de studie van hoofdstuk 2 is dat met zekerheid bewezen kan worden of een bepaalde akoestische parameter gebruikt wordt voor discriminatie of niet. Aan de andere kant kunnen met deze methode slechts twee parameters tegelijk getest worden, wat het testen van grote aantallen parameters tot een langdurende klus zou maken. Daarbij is het gebruik van synthetische stimuli een nadeel, omdat we uiteindelijk geïnteresseerd zijn in de perceptie van echte koervocalisaties.

In hoofdstuk 3 werd een methode gebruikt die minder direct is in de identificatie van perceptueel belangrijke koereigenschappen, maar die als voordeel

heeft dat veel parameters tegelijkertijd getest kunnen worden. Bovendien kan gebruik gemaakt worden van natuurlijke koergeluiden. In dit experiment werden duiven getraind om alleen te reageren op hun soortspecifieke koergeluid. Vervolgens kregen ze de koergeluiden van 12 andere *Streptopelia* soorten (allospecifieke koeren) te horen, en de sterkte van de respons werd gebruikt als maat voor hun gelijkenis op de soortspecifieke geluiden. De resultaten laten zien dat de duiven onderscheid maken tussen de allospecifieke koergeluiden: sommige lijken meer op hun eigen koergeluid dan andere. Met statistische technieken werden drie akoestische parameters geïdentificeerd die correleren met perceptuele gelijkenis: koerduur, minimum frequentie, en Wiener entropie. Het relatieve belang van deze parameters wordt bepaald door hun percipieerbaarheid en door de mate waarin ze variëren tussen koergeluiden. Koerduur blijkt verreweg de belangrijkste parameter bij het onderscheiden van soortspecifieke van allospecifieke koergeluiden. Bovendien blijkt de gevoeligheid voor verschillen in duur groter dan verwacht op basis van eerdere gehoorsstudies met simpele, korte tonen. Dit onderstreept dat het perceptuele belang van vocalisatiekarakteristieken het best onderzocht kan worden met behulp van echte vocalisaties, of synthetische stimuli die in voldoende mate lijken op echte vocalisaties.

Uit een vergelijking van de resultaten in hoofdstuk 3 met die van een eerdere studie naar de afstamming van *Streptopelia* soorten, blijkt dat er geen duidelijke relatie bestaat tussen de genetische afstand tussen *Streptopelia* soorten en de perceptuele gelijkenis van hun perch-koeren. Dit is een verassend resultaat omdat culturele evolutie geen rol speelt in de evolutie van koergeluiden. Evolutionaire veranderingen in de geïdentificeerde parameters vinden daarom plaats door middel van natuurlijke en seksuele selectie, en genetische 'drift'. De snelheid van evolutionaire verandering die door zulke processen veroorzaakt wordt is minder groot dan wanneer er ook sprake is van culturele evolutie. Desondanks is de evolutie van koervocalisaties blijkbaar snel genoeg om een duidelijk afstammingspatroon in hun perceptuele gelijkenis teniet te doen.

In het tweede deel van dit proefschrift, hoofdstuk 4 en 5, beschrijf ik twee studies naar geluidsproductiemechanismen bij *Streptopelia* duiven. De experimentele technieken in beide studies zijn gelijk: in spontaan vocaliserende vogels werden de patronen van luchtstroom door de luchtpijp en luchtdruk in verschillende luchtzakken geregistreerd. Het doel van de twee studies was echter zeer verschillend.

Met de studie in **hoofdstuk 4** werd onderzocht hoe *Streptopelia* duiven in staat zijn om puur-tonale geluiden te produceren. De meeste *Streptopelia* koeren zijn puur-tonaal, maar sommige zijn dat niet. Er is een langlopende

discussie in de vogelzangliteratuur over de vraag of puur-tonale vogelzang verklaard kan worden met een geluidsproducerend mechanisme dat in principe werkt als een pulserende klep. Zo'n klepmechanisme is de normale manier van vocale productie bij andere dieren, maar leidt in de regel tot de vorming van sterke harmonische boventonen in het geluid. In puur-tonale vogelzang ontbreken zulke boventonen. In deze studie werden geluidsignalen onderzocht die aanwezig zijn in de luchtpijp en luchtzakken. Hieruit blijkt dat het koergeluid binnen in het dier, dicht bij het geluidsproducerend orgaan, wel harmonische boventonen bevat. Dit laat zien dat het niet nodig is om twee fundamenteel verschillende bronmechanismen aan te wenden, om het verschil tussen puur-tonale koeren en koeren met een complexere tonale structuur te verklaren. Puur-tonaal geluid wordt blijkbaar gegenereerd door filtering van het brongeluid voordat het geluid radieert van het dier. Deze vondst vormt tot nog toe het meest directe bewijs voor het bestaan van een bronfilter mechanisme in de productie van vogelzang, en is ook van belang buiten de context van vogelzang. Het bestaan van een multi-frequentie, harmonisch brongeluid dat gefilterd wordt, betekent een uitbreiding van de al bekende overeenkomsten tussen vogelzang en menselijk spraak met het mechanisme van akoestische productie.

In hoofdstuk 5 onderzoek ik hoe frequentie en amplitude gemoduleerd worden in de vocalisaties van *S. risoria*. Uit de patronen van luchtzakdruk en tracheale luchtstroom tijdens vocalisatie blijkt dat het tritype amplitudemodulatie veroorzaakt wordt door het cyclisch openen en sluiten van een klep. Dit in tegenstelling tot de tot nu toe gangbare hypothese van oscillerende luchtzakdrukken. De patronen van frequentiemodulatie blijken veel complexer dan tot nog toe werd herkend. Ze bestaan uit graduele, continue tijd-frequentiepatronen, die worden onderbroken door plotselinge, discontinue frequentiesprongen. Het patroon van graduele frequentiemodulatie blijkt zeer nauw te corresponderen met het luchtdrukpatroon in de interclaviculaire luchtzak. Ik beargumenteer dat de druk in de interclaviculaire luchtzak hoogstwaarschijnlijk de veroorzakende factor is van frequentiemodulatie. Hoe frequentiesprongen worden veroorzaakt is nog onbekend, maar een waarschijnlijke verklaring is dat ze te maken hebben met intrinsieke, niet-lineaire eigenschappen van het geluidsproducerend orgaan. Luchtzakdrukvariatie als mechanisme voor frequentiemodulatie is een nieuwe toevoeging aan de bekende manieren waarop vogels de toonhoogte van hun zang reguleren. Dit mechanisme contrasteert met de gespecialiseerde spieren die zangvogels hiervoor gebruiken, en kan verklaren waarom de toonhoogte in niet-zangvogels in het algemeen maar in beperkte mate gemoduleerd wordt.

Verder is interessant dat de resultaten van hoofdstuk 5 suggereren dat er

maar twee centraal gecoördineerde motorvariabelen betrokken zijn bij het genereren van modulatiecomplexiteit in de koeren van *S. risoria*. Simpele en geleidelijke luchtdrukpatronen bepalen de hoofdpatronen van temporele structuur en graduele frequentiemodulatie, terwijl amplitudemodulatie veroorzaakt wordt door het cyclisch openen van een klep. Veel van de rest van de complexiteit – oscillatoire frequentiemodulatie en frequentiesprongen – kan verklaard worden met mechanismen die geen centrale controle vergen: feedback en intrinsieke, niet-lineaire eigenschappen van het geluidsproducerend orgaan.

Conclusie

Mijn proefschrift laat zien dat hoewel temporele structuur alleen al voldoende is om *Streptopelia* koeren naar soort te kunnen onderscheiden, duiven toch meerdere akoestische parameters gebruiken om het onderscheid te maken. Ik heb de belangrijkste parameters hiervan geïdentificeerd. Dit verschaft niet alleen een beter inzicht in hoe duiven vocalisaties waarnemen, maar geeft ook aan hoe koergeluiden het meest efficiënt zouden kunnen divergeren bij soortvorming, althans perceptueel gezien. De vondst dat verschillen in een temporele parameter voor duiven belangrijker zijn dan voor andere vogels, wijst erop dat perceptuele mechanismen geëvolueerd zijn in reactie op de grote soortverschillen in deze parameter. Hoewel dit argument eigenlijk ook omgedraaid zou kunnen worden, laat het zien dat mechanismen en evolutie van communicatie met elkaar te maken hebben en niet afzonderlijk van elkaar bestudeerd moeten worden.

Vervolgens heb ik laten zien welke perifere mechanismen betrokken zijn bij het genereren van specifieke modulatiepatronen in koergeluiden, en dus verantwoordelijk zijn voor een deel van de akoestische verschillen tussen *Streptopelia* soorten. Met de identificatie van interclaviculaire luchtzakdruk als causale factor in frequentiemodulatie, lever ik een verklaring voor waarom frequentiemodulatie in duiven – en wellicht ook in andere niet-zangvogels – beperkt is vergeleken met zangvogels. Deze mechanische beperking in frequentiemodulatie verklaart op haar beurt weer waarom soortverschillen in koergeluiden vooral aanwezig zijn in temporele parameters. Het grote gewicht dat door perceptiemechanismen toegekend wordt aan een temporele parameter heeft daarom niet alleen te maken met de evolutie van soortspecifieke vocalisaties, maar ook met de onderliggende productiemechanismen.

References

- Ballintijn, M. R. (1999). *Vocal variation in the collared dove: coding messages in coo-vocalizations*. PhD thesis, Leiden University, The Netherlands.
- Ballintijn, M. R. and ten Cate, C. (1998). Sound production in the collared dove: a test of the ‘whistle’ hypothesis. *Journal of Experimental Biology*, 201:1637–1649.
- Banta Lavanex, P. (1999). Vocal production mechanisms in the budgerigar (*Melopsittacus undulatus*): The presence and implications of amplitude modulation. *Journal of the Acoustical Society of America*, 106:491–505.
- Baptista, L., Trail, P. W., and Horblit, H. M. (1997). Family Columbidae (pigeons and doves). In del Hoyo, J., Elliot, A., and Sargatal, J., editors, *Sandgrouse to Cuckoos*, volume 4 of *Handbook of the Birds of the World*, pages 60–243. Lynx Edicions, Barcelona.
- Becker, P. H. (1982). The coding of species-specific characteristics in bird sounds. In Kroodsma, D. E. and Miller, E. H., editors, *Acoustic Communication in Birds*, volume 1, chapter 7, pages 213–252. Academic Press, New York.
- Beckers, G. J. L., Goossens, M. A., and ten Cate, C. (2002a). Perceptual salience of acoustic differences between conspecific and allospecific vocalizations in turtle-doves. *Animal Behaviour*. *In press*.
- Beckers, G. J. L., Suthers, R. A., and ten Cate, C. (2002b). Pure-tone birdsong by resonance filtering of harmonic overtones. *Submitted*.
- Beckers, G. J. L. and ten Cate, C. (2001). Perceptual relevance of species-specific differences in acoustic signal structure in *Streptopelia* doves. *Animal Behaviour*, 62:511–518.

- Bennet-Clark, H. C. (1999). Resonators in insect sound production: how insects produce loud pure-tone songs. *Journal of Experimental Biology*, 202:3347–3357.
- Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences*, 17:97–110.
- Brackenbury, J. H. (1977). Physiological energetics of cock-crow. *Nature*, 270:433–435.
- Brackenbury, J. H. (1980). Control of sound production in the syrinx of the fowl *Gallus gallus*. *Journal of Experimental Biology*, 85:239–251.
- Bradbury, J. W. and Vehrencamp, S. L. (1998). *Principles of animal communication*. Sinauer Associates, Sunderland, Massachusetts.
- Brittan-Powell, E. F., Dooling, R. J., Larsen, O. N., and Heaton, J. T. (1997). Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*). *Journal of the Acoustical Society of America*, 101:578–589.
- Brown, S. D., Dooling, R. J., and O’Grady, K. (1988). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): III. Contact calls. *Journal of Comparative Psychology*, 102:236–247.
- Casey, R. M. and Gaunt, A. S. (1985). Theoretical models of the avian syrinx. *Journal of Theoretical Biology*, 116:45–64.
- Cheng, M. F., Peng, J. P., and Johnson, P. (1998). Hypothalamic neurons preferentially respond to female nest coo stimulation: demonstration of direct acoustic stimulation of luteinizing hormone release. *Journal of Neuroscience*, 18:5477–5489.
- Cuvier, G. and Duvernoy, G. L. (1846). *Leçons d’anatomie comparée*, volume 18. Paris: Fortin, Masson et Cie, 2nd edition.
- Cynx, J. (1995). Similarities in absolute and relative pitch perception in songbirds (starling and zebra finch) and a nonsongbird (pigeon). *Journal of Comparative Psychology*, 109:261–267.
- Cynx, J. and Clark, S. J. (1998). The laboratory use of conditional and natural responses in the study of avian auditory perception. In Hopp, S. L., Owren, M. J., and Evans, C. S., editors, *Animal acoustic communication*, pages 353–377. Springer, Berlin.

- de Kort, S. R. (2002). *Pigeon pidgin: evolution of vocal signals through interspecific interactions in Streptopelia doves*. PhD thesis, Leiden University, The Netherlands.
- de Kort, S. R. and ten Cate, C. (2001). Response to interspecific vocalizations is affected by degree of phylogenetic relatedness in *Streptopelia* doves. *Animal Behaviour*, 61:239–247.
- de Kort, S. R. and ten Cate, C. (2002). When fewer is better: repeated decrease in vocal repertoire size in non-songbirds, *Streptopelia* doves. In preparation.
- Dooling, R. J. (1982). Auditory perception in birds. In Kroodsma, D. E. and Miller, E. H., editors, *Acoustic communication in birds*, volume 1, chapter 4, pages 95–130. Academic Press, New York.
- Dooling, R. J., Brown, S. D., Klump, G. M., and Okanoya, K. (1992). Auditory perception of conspecific and heterospecific vocalizations in birds: Evidence for special processes. *Journal of Comparative Psychology*, 106:20–28.
- Dooling, R. J., Brown, S. D., Park, T. J., and Okanoya, K. (1987a). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): I. Pure tones. *Journal of Comparative Psychology*, 101:139–149.
- Dooling, R. J., Park, T. J., Brown, S. D., Okanoya, K., and Soli, S. D. (1987b). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): II. Vocal signals. *Journal of Comparative Psychology*, 101:367–381.
- Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22:567–631.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
- Fee, M. S., Shraiman, B., Pesaran, B., and Mitra, P. P. (1998). The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature*, 395:67–71.
- Flanagan, J. L. (1972). *Speech analysis, synthesis and perception*. Berlin: Springer, 4th edition.
- Fletcher, N. H. (1989). Acoustics of bird song – some unresolved problems. *Comments on Theoretical Biology*, 1:237–251.

- Fletcher, N. H. (1992). *Acoustic systems in biology*. Oxford: Oxford University Press.
- Fletcher, N. H. (2000). A class of chaotic bird calls? *Journal of the Acoustical Society of America*, 108:821–826.
- Fletcher, N. H. and Tarnopolsky, A. (1999). Acoustics of the avian vocal tract. *Journal of the Acoustical Society of America*, 105:35–49.
- Gardner, T., Cecchi, G., Magnasco, M., Laje, R., and Mindlin, G. B. (2001). Simple motor gestures for birdsongs. *Physical Review Letters*, 87:art. no. 208101.
- Gaunt, A. S. (1983). A hypothesis concerning the relationship of syringeal structure to vocal abilities. *Auk*, 100:853–862.
- Gaunt, A. S. (1986). Interaction of syringeal structure and airflow in avian phonation. In Ouellet, H., editor, *XIX Congressus Internationalis Ornithologici*, pages 913–924. University of Ottawa Press, Ottawa.
- Gaunt, A. S., Gaunt, S. L. L., and Casey, R. M. (1982). Syringeal mechanisms reassessed: evidence from *Streptopelia*. *Auk*, 99:474–494.
- Gaunt, A. S., Gaunt, S. L. L., and Hector, D. H. (1976). Mechanics of the syrinx in *Gallus gallus*. I A comparison of pressure events in chickens to those in oscines. *The Condor*, 78:208–223.
- Gaunt, A. S., Gaunt, S. L. L., and Hector, D. H. (1977). Mechanics of the syrinx in *Gallus gallus*. II Electromyographic studies of ad libitum vocalizations. *Journal of Morphology*, 152:1–20.
- Gaunt, A. S. and Nowicki, S. (1998). Sound production in birds: acoustics and physiology revisited. In Hopp, S. L., Owren, M. J., and Evans, C. S., editors, *Animal acoustic communication*, chapter 4, pages 291–321. Springer, Berlin.
- Genter, T. Q. and Hulse, S. H. (2000). Perceptual classification based on the component structure of song in European starlings. *Journal of the Acoustical Society of America*, 107:3369–3381.
- Gerhardt, H. C. (1991). Female mate choice in treefrogs: Static and dynamic acoustic criteria. *Animal Behaviour*, 42:615–635.

- Goller, F. and Larsen, O. N. (1997a). In situ biomechanics of the syrinx and sound generation in pigeons. *Journal of Experimental Biology*, 200:2165–2176.
- Goller, F. and Larsen, O. N. (1997b). A new mechanism of sound generation in songbirds. *Proceeding of the National Academy of Sciences USA*, 94:14787–14791.
- Goller, F. and Suthers, R. A. (1996). Role of syringeal muscles in controlling the phonology of bird song. *Journal of Neurophysiology*, 76:287–299.
- Goodwin, D. (1983). *Pigeons and doves of the world*. British Museum of Natural History, London.
- Greenwalt, C. H. (1968). *Bird song: acoustics and physiology*. Washington: Smithsonian Institution Press.
- Gürtler, W. (1973). Artisolierende Parameter der Rivierrufs der Türkentaube (*Streptopelia decaocto*). *Journal für Ornithologie*, 114:305–316.
- Hoese, W. J., Podos, J., Boetticher, N. C., and Nowicki, S. (2000). Vocal tract function in birdsong production: experimental manipulation of beak movements. *Journal of Experimental Biology*, 203:1845–1855.
- Hulse, S. H. (1995). The discrimination-transfer procedure for studying auditory perception and perceptual invariance in animals. In Klump, G., Dooling, R., Fay, R., and Stebbins, W., editors, *Methods in comparative psychoacoustics*, volume 6, pages 319–330. Birkhäuser Verlag, Basel.
- Iverson, P. and Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, 97:553–562.
- Johnson, K. P., de Kort, S., Dinwoodey, K., Mateman, A. C., ten Cate, C., Lessells, C. M., and Clayton, D. H. (2001). A molecular phylogeny of the dove genus *Streptopelia*. *Auk*, 118:874–887.
- Johnsrude, I. S., Weary, D. M., Ratcliffe, L. M., and Weisman, R. G. (1994). Effect of motivational context on conspecific song discrimination by brown-headed cowbirds (*Molothrus ater*). *Journal of Comparative Psychology*, 108:172–178.

- King, A. S. (1989). Functional anatomy of the syrinx. In King, A. S. and McLelland, J., editors, *Form and function in birds*, volume 4, pages 105–192. Academic Press, London.
- Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (2000). *Fundamentals of acoustics*. New York: John Wiley & Sons, 4th edition.
- Lade, B. I. and Thorpe, W. H. (1964). Dove songs as innately coded patterns of specific behaviour. *Nature*, 202:366–368.
- Larsen, O. N. and Goller, F. (1999). Role of syringeal vibrations in birds vocalizations. *Proceedings of the Royal Society London B*, 266:1609–1615.
- Larsen, O. N. and Goller, F. (2002). Direct observation of syringeal muscle function in songbirds and a parrot. *Journal of Experimental Biology*, 205:25–35.
- Lehrman, D. S. (1965). Interaction between internal and external environments in the regulation of the reproductive cycle of the ring dove. In Beach, F. A., editor, *Sex and behaviour*, pages 355–380. Wiley, New York.
- LePrell, C. G. and Moody, D. B. (1997). Perceptual salience of acoustic features of japanese monkey coo calls. *Journal of Comparative Psychology*, 111:261–274.
- Lockner, F. R. and Murrish, D. E. (1975). Interclavicular air sac pressures and vocalization in mallard ducks *Anas platyrhynchos*. *Comparative Biochemistry and Physiology A*, 52:183–187.
- Mbu Nyamsi, R. G., Aubin, T., and Bremond, J. C. (1994). On the extraction of some time dependent parameters of an acoustic signal by means of the analytic signal concept. Its application to animal sound study. *Bioacoustics*, 5:187–203.
- Michelsen, A. (1983). Biophysical basis of sound communication. In Lewis, B., editor, *Bioacoustics, a comparative approach*, chapter 1, pages 3–38. Academic Press, London.
- Miller, E. H. (1982). Character and variance shift in acoustic signals of birds. In Kroodsma, D. E. and Miller, E. H., editors, *Acoustic Communication in Birds*, volume 1, chapter 8, pages 253–295. Academic Press, New York.
- Miller, W. J. and Miller, L. S. (1958). Synopsis of behaviour traits of the ring neck dove. *Animal Behaviour*, 6:3–8.

- Miskimen, M. (1951). Sound production in passerine birds. *Auk*, 68:493–504.
- Moore, B. C. J. (1997). *An introduction to the psychology of hearing*. Academic Press, San Diego, 4th edition.
- Nottebohm, F. (1976). Phonation in the orange-winged Amazon parrot, *Amazona amazonica*. *Journal of Comparative Physiology*, 108A:157–170.
- Nottebohm, F. and Nottebohm, M. E. (1971). Vocalizations and breeding behavior of surgically deafened Ring Doves, *Streptopelia risoria*. *Animal Behaviour*, 19:313–327.
- Nowicki, S. (1987). Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature*, 235:53–55.
- Nowicki, S. and Marler, P. (1988). How do birds sing? *Music Perception*, 5:391–426.
- Okanoya, K. and Dooling, R. J. (1991). Perception of distance calls by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Poephila guttata*): Assessing species-specific advantages. *Journal of Comparative Psychology*, 105:60–72.
- Okanoya, K. and Kimura, T. (1993). Acoustical and perceptual structures of sexually dimorphic distance calls in bengalese finches (*Lonchura striata domestica*). *Journal of Comparative Psychology*, 107:386–394.
- Podos, J. (1997). A performance constraint on the evolution of trilled vocalizations in a songbird family (passeriformes: Emberizidae). *Evolution*, 51:537–551.
- Rüppell, W. (1933). Physiologie und Akustik der Vogelstimme. *Journal für Ornithologie*, 81:433–542.
- Simmons, A. M. and Bean, M. E. (2000). Perception of mistuned harmonics in complex sounds by the bullfrog (*Rana catesbeiana*). *Journal of Comparative Psychology*, 114:167–173.
- Sinnot, J. M. (1980). Species-specific coding in bird song. *Journal of the Acoustical Society of America*, 68:494–497.
- Sinnot, J. M. and Saporita, T. A. (2000). Differences in American English, Spanish, and monkey perception of the say-stay trading relation. *Perception and Psychophysics*, 62:1312–1319.

- Slabbekoorn, H. (1998). *Perception of coo variation by the collared dove: a parameter analysis using playbacks in the field*. PhD thesis, Leiden University, The Netherlands.
- Slabbekoorn, H., de Kort, S., and ten Cate, C. (1999). Comparative analysis of perch-coo vocalizations in *Streptopelia* doves. *Auk*, 116:737–748.
- Slabbekoorn, H. and ten Cate, C. (1998). Perceptual tuning to frequency characteristics of territorial signals in collared doves. *Animal Behaviour*, 55:847–857.
- Slabbekoorn, H. and ten Cate, C. (1999). Collared dove responses to playback: Slaves to the rhythm. *Ethology*, 105:377–391.
- Sokal, R. R. and Rohlf, F. J. (1995). *Biometry: the principles and practice of statistics in biological research*. W.H. Freeman and Company, New York, 3rd edition.
- Stein, R. C. (1973). Sound production in vertebrates: summary and prospectus. *American Zoologist*, 13:1249–1255.
- Suthers, R. A. and Goller, F. (1997). Motor correlates of vocal diversity in songbirds. In Nolan Jr, V., Ketterson, E., and Thompson, S. F., editors, *Current Ornithology*, volume 14, chapter 5, pages 235–288. Plenum Press, New York.
- Suthers, R. A., Goller, F., and Hartley, R. S. (1994). Motor dynamics of song production by mimic thrushes. *Journal of Neurobiology*, 25:917–936.
- Suthers, R. A., Goller, F., and Pytte, C. (1999). The neuromuscular control of birdsong. *Philosophical Transactions of the Royal Society B*, 354:927–939.
- Suthers, R. A., Goller, F., and Wild, J. M. (2002). Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proceedings of the National Academy of Sciences USA*, 99:5680–5685.
- Suthers, R. A. and Hector, D. H. (1985). The physiology of vocalization by the echolocating oilbird, *Steatornis caripensis*. *Journal of Comparative Physiology A*, 156:243–266.
- Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B., and Mitra, P. P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour*, 59:1167–1176.

- ten Cate, C. (1992). Coo types in the collared dove *Streptopelia decaocto*: one theme, distinctive variations. *Bioacoustics*, 4:161–183.
- ten Cate, C., Slabbekoorn, H., and Ballintijn, M. R. (2002). Birdsong and male–male competition: Causes and consequences of vocal variability in the collared dove (*Streptopelia decaocto*). *Advances in the Study of Behaviour*, 31:31–75.
- Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie*, 20:410–433.
- Weary, D. M. (1990). Categorization of song notes in great tits: Which acoustic features are used and why? *Animal Behaviour*, 39:450–457.
- Wilden, I., Herzog, H., Peters, G., and Tembrock, G. (1998). Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics*, 9:171–196.
- Wiley, R. H. and Richards, D. G. (1982). Adaptation for acoustic communication in birds: Sound transmission and signal detection. In Kroodsma, D. E. and Miller, E. H., editors, *Acoustic Communication in Birds*, volume 1, chapter 5, pages 131–181. Academic Press, New York.
- Wytenbach, R. A., May, M. L., and Hoy, R. R. (1996). Categorical perception of sound frequency by crickets. *Science*, 273:1542–1544.
- Youngren, O. M., Peek, F. W., and Phillips, R. E. (1974). Repetitive vocalizations evoked by local electrical stimulation of avian brains. III Evoked activity in the tracheal muscles of the chicken (*Gallus gallus*). *Brain, Behavior and Evolution*, 9:393–421.

Acknowledgements

I would like to acknowledge those who provided support, one way or the other, during my dissertation work.

Rod Suthers invited me to carry out part of the vocal production work at his lab at Indiana University, USA, and has provided invaluable guidance and support. Our cooperative effort was the basis of the last two chapters in this thesis. The psychoacoustic work has been carried out with the help of Aldwin van Broekhoven, Renske Gilissen, Brigitte Goossens and Sebastiaan Ruinard, who participated in this research as part of their undergraduate program. I hope the benefit was mutual. Sandy Ronan and Peter Snelderwaard helped with experimental work on vocal production physiology. Koen Verhoeven provided me with advice and discussion on statistical methodology. Albertina Leitão and Nienke Terpstra proof-read this thesis.

Over the years I had interesting discussions about science and other matters with my Leiden colleagues Selvino de Kort, Paula den Hartog, Eric Etman, Bart Houx, Rob Lachlan, Albertina Leitão, Katharina Riebel, Jean Secondi, Hans Slabbekoorn, Nienke Terpstra, and Machteld Verzijden. Some of these discussions found their way into this thesis. I also appreciate the support from my other Leiden colleagues, Enja Feuth, Gerdi Dijkstra, and Hans Dudart. I had fruitful discussions specifically on bird song production with Coen Elemans. Brian Nelson and Sue Anne Zollinger improved the quality of my work and life (these are not always separable in science) during my lab visit in Bloomington. Furthermore, I very much appreciate that Brigitte Goossens and Albertina Leitão were willing to be my ‘paranimfen’.

Finally I want to thank my girlfriend Marjolein Efting Dijkstra for putting up with all the burdens associated with my work. She has been a continuous source of support and motivation throughout the whole project.

Curriculum vitae

I was born 10 March 1970 in Urmond, the Netherlands. In 1988 I completed Gymnasium β at the Serviam Lyceum in Sittard, and started to study Biology at the university of Nijmegen. After a specialization in physiology and molecular biology, I carried out a nine-month undergraduate research project on brain activity imaging in the clawed toad *Xenopus laevis*, at the group of Cellular Animal Physiology, supervised by Prof. dr Eric Roubos. I was introduced to biology outside the lab in a nine-month additional project on warning colouration and mimicry in coral snakes (genus *Micrurus*), in the rainforest reserve of Rara Avis, Costa Rica, Central America, supervised by dr Donald Perry and dr Henk Strijbosch. Autumn 1994 I received a MSc degree in biology.

During my undergraduate work in Costa Rica I became impressed with the Costa-Rican rainforest biodiversity, and worried by its rapid disappearance. After graduation, I therefore employed my scientific training for applied research projects into rainforest conservation. Together with colleague biologist and friend Koen Verhoeven, I set up Bioproca Foundation, a nonprofit organization for the development of innovative methods that contribute to long-term conservation of tropical rainforests. We focussed on canopy orchids, and carried out the practical part of the work in Rara Avis.

Early 1998 I switched back to academic science, and took on a PhD project on mechanisms of acoustic communication in *Streptopelia* doves, at the Behavioural Biology group of Prof. dr Carel ten Cate, Institute of Evolutionary and Ecological Sciences at Leiden University. Part of the results of this work are reported in this thesis.

Since October 2002 I continue my work on vocal production physiology in birds, now as a postdoctoral fellow at the laboratory of Prof. dr Roderick Suthers, School of Medicine, Indiana University, USA.

