

Universiteit
Leiden
The Netherlands

The music of language : exploring grammar, prosody and rhythm perception in zebra finches and budgerigars

Spierings, M.J.

Citation
Spierings, M. J. (2016, November 17). *The music of language : exploring grammar, prosody and rhythm perception in zebra finches and budgerigars*. Retrieved from https://hdl.handle.net/1887/44267

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/44267

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle http://hdl.handle.net/1887/44267 holds various files of this Leiden University dissertation.

Author: Spierings, M.J.
Title: The music of language : exploring grammar, prosody and rhythm perception in zebra finches and budgerigars
Issue Date: 2016-11-17

The Music of Language

Exploring prosody, rhythm and grammar perception
in zebra finches and budgerigars

Michelle Spierings

Spierings, Michelle Johanna

The Music of Language
Exploring prosody, rhythm and grammar perception in zebra finches and budgerigars

Dissertation Leiden University

An electronic version of this thesis can be downloaded from:
openacces.leidenuniv.nl

Printed by Ridderprint, Ridderkerk
Artwork by Andrea van den Berg
Edited by Michelle Spierings

© 2016

The music of language

Exploring prosody, rhythm and grammar perception
in zebra finches and budgerigars.

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan Universiteit Leiden,
op gezag van Rector Magnificus Prof. Mr. C.J.J.M. Stolker
volgens besluit van het College van Promoties
te verdedigen op 17 november 2016
klokke 13:45 uur

door

Michelle Johanna Spierings

geboren op 18 juli 1987
te Voorburg, Nederland

Promotiecommissie

Promotor

Prof. dr. C. J. ten Cate

Overige leden

Prof. dr. W. T. Fitch, University of Vienna

Prof. dr. C. Scharff, Freie Universität Berlin

Dr. G. J. L. Beckers, Utrecht University

Prof. dr. A. H. Meijer

Prof. dr. H. P. Spaink

Dit onderzoek werd gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), 360.70.452.

Table of contents

Chapter 1	General introduction	**9**
Chapter 2	Pauses enhance chunk recognition in song element strings by zebra finches.	**34**
Chapter 3	Zebra finches group tones alternating in pitch as trochees.	**54**
Chapter 4	Zebra finches are sensitive to prosodic features of human speech.	**72**
Chapter 5	Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species.	**92**
Chapter 6	Zebra finches as a model species to understand the roots of rhythm.	**130**
Chapter 7	Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment.	**140**
Chapter 8	Thesis summary & General discussion	**171**
	Nederlandse samenvatting	**189**
	Acknowledgements	**195**
	Curriculum Vitae	**199**
	Publications	**203**



1

GENERAL INTRODUCTION



GENERAL INTRODUCTION

Language is a complex cognitive system of rules and regularities that enables us to communicate about our thoughts, experiences and intentions. It is considered to be a uniquely human trait. All animals have ways to communicate, but these communication systems do not seem to reach the semantic and syntactic complexity of human language. It is hotly debated how the human language faculty arose. What is the core of language and which mechanisms form the bedrock of the human language faculty? If there was some sort of protolanguage, how was it structured? These are a few of the main questions in the debate on language evolution and the answers differ wildly. There are scholars hypothesizing that the ability to use complex syntax appeared rather sudden in the course of evolution, creating a special language ability only in humans. This rather sudden start of language is ascribed to the evolution of universal grammar, a language proficiency shared between all humans (e.g. 1,2). Others argue that a unification of gesture and speech was at the root of language development. As thoughts can be expressed simultaneously in both speech and gesture, and speech and gesture are integrated at a neural level, they might have shaped our language ability in interaction (3). Furthermore, it is debated which selection pressures were involved in the evolution of language. Language might have evolved by the strong pressures of natural selection alone (4) or by a combination of natural selection, pre-adaptations and cultural transmission with learning (5).

It is clear that the theories on language evolution vary extensively, but they share one thing: the lack of empirical evidence. Language did not leave archeological traces until humanity developed writing, which was, unarguably, much later than the emergence of language. This means that we only have our current knowledge to try and decipher what might have been at the root of language evolution. In this thesis I will add to the discussion on the roots of language evolution with comparative studies on language and music perception in two bird species. Similar cognitive mechanisms in distantly related species can give us insight in the abilities that might have been present in early hominids.

The language faculty and language evolution

Producing and perceiving language is not based on a single cognitive ability. On the contrary, it is a set of different abilities or mechanisms that together form the language faculty (6). One of the first crucial aspects of language learning is the ability to adjust one's own vocalizations to match the vocalizations produced by others. This is called vocal learning and is thus far only known to be present in humans and a few

animal classes (cetaceans (7), pinnipeds (8-10), bats (11), elephants (12), hummingbirds (13), parrots (14), and songbirds (15)). Furthermore, in order to produce or perceive language, one needs to be able to produce different speech sounds and hear the difference between them, an ability that allows differentiation between spoken words. Finally, listeners need to be able to grasp the syntactic structure of a sentence to understand the difference between questions and statements or to determine the function of a particular word. These and many other cognitive perceptual mechanisms shape the language faculty and together allow us to communicate the way we do.

In the discussion on mechanisms involved in language and which mechanisms are uniquely human, Hauser, Fitch and Chomsky proposed a division of mechanisms in the "Faculty of language in the narrow sense" (FLN) and the "Faculty of language in the broad sense" (FLB) (16,17). The FLB includes all mechanisms required for language, such as perceiving speech, learning grammar and understanding intentions. The FLN includes only those mechanisms that are unique to human language. Although the division of mechanisms in the FLN and FLB might be a point of discussion and will likely change with more empirical data, it does help us to consider the aspects of language that might or might not be unique to language. The shared mechanisms most likely did not evolve specifically for language. They might have been present in a pre-linguistic situation, providing the basis for a possible proto-language.

Speech and Language

Speech is the production of sound units used to express language. Spoken language is the product of specific combinations of these speech sounds into sentences that convey meaning. In order to correctly perceive speech, the listener needs to comprehend the difference between the sound units, phonemes. The complexity of phoneme discrimination is the categorization of phonemes produced by different speakers. Vowels are, for example, distinguishable by the relative distance between the prominent frequency bands in the sound, also called the formants. An /e/ produced by a male has formants in different positions and has a lower fundamental frequency than an /e/ produced by a female speaker. Nevertheless, 3-6 month old infants are already able to categorize these speech sounds as belonging to the same vowel category (e.g. 18,19). This has long been thought to be unique to humans (20,21) but we now know that animal species are also able to categorize phonemes (22-24).

In order to communicate, phonemes need to be combined into meaningful words and sentences in a, for the language of the speaker, grammatically correct order. Similar to speech sounds, listeners can discriminate different phoneme orders as forming the

words of the language. Already from 7 months old, infants can detect the order of phonemes and generalize this order to new speech sounds (25). Whether non-human animals can generalize language patterns is still debated.

Phonological syntax versus lexical syntax

Syntax is the particular structure in which sounds are ordered to form words, sentences and phrases. Both speech and language have a particular syntax: known as phonological and lexical syntax. The phonological regularities in a language describe which combinations of sounds are likely to form words and which ones are unlikely to be words. For example, in English, the /ŋ/ at the end of "sing" never occurs at the beginning of a word and the /h/ in "hat" never occurs at the end of a word (26). Even more complex structures, such as non-adjacent dependencies, follow phonological syntactic rules. In these cases, a sound at the beginning of a word determines which sound is likely to be at the end of the same word (27,28). Besides a likely order of sounds, languages also have a grammatically correct order of nouns, vowels and adjectives, the compositional syntax. For example, the sentences "dog bites man" and "man bites dog" are both grammatically correct. Nevertheless, "dog bites man" is a more likely utterance. Infants can learn to distinguish different compositions of items already during early language acquisition (29,30).

The distinction between phonological and lexical syntax is important for studies as the ones presented in this thesis. The computational complexity of a task is related to the type of syntax. In phonological syntax learning, the subject learns the difference between the orders in which phonemes can occur. When learning lexical syntax, however, the subject has to first form sound categories. It has to learn that a group of words belong to one specific category (e.g. nouns) and other words fall into another category (e.g. verbs). Both abilities are often tested in artificial grammar learning paradigms where infants are exposed to an artificially constructed miniature language and are later tested on whether they acquired the syntax of this language (e.g. 31). In a review by Gomez and Gerken (32), they set the two types of syntax learning against each other and concluded that there are different computations at the basis of each syntax learning ability.

Language and music evolution

Language is not the only complex universal cognitive trait that seems to be uniquely human. Our ability to produce and synchronize with musical structures bears the same fascinating sense of complexity and universality as language. Music is produced by humans of all cultures and, across cultures, infants already have an increased interest

in music over other non-human sounds (33,34). The function of music has been much debated, ranging from music being used for and developed through sexual selection (35), to a function of music for bonding between parents and infants or between other individuals in a social group (36,37). The oldest musical instrument found dates from 45 000 years ago, a short time when considering the evolution of humans. Rhythmic abilities likely arose earlier than the production of musical instruments, but the lack of fossil artifacts makes a definite determination of the function and evolution of music a daunting task (38).

Regardless of the clear differences between music and language, there are also analogies between the two. Both have a temporal structure, with the information arriving in a particular order over time. Furthermore, music is structured according to a particular rhythm that can be generalized to new tones or faster and slower versions. Languages are also distinguishable by their typical rhythmic structures and also in language the rhythm can be generalized to new words. Finally, each language has a certain prosody, the pitch contour of words and the prosodic pattern of sentences. Similar sound features can be found in music.

Given the similarities between music and language, it is relevant to study the mechanisms involved in both language and music perception. There are two main theories regarding the root of the music ability (39,40). The more classical view on music evolution is that it is a by-product of our ability to speak and developed as a trade-off of language, an epiphenomenon. In this view, music is subject to culture with no evolution of its own (41). A more recent view is that musical abilities might have evolved separately, independent from language and as a specific cognitive adaptation (38). Empirical data is needed to shed more light on the underlying mechanisms of both language and music, not only to determine their current influences on each other, but also to consider how these two systems might have evolved.

Comparative studies

One important method to study the evolution of language is by studying which cognitive mechanisms of the language faculty are uniquely human and which mechanisms are shared with other species (e.g. 16). Shared cognitive abilities are unlikely to have evolved specifically for language and might have been present in our far, pre-linguistic, ancestors. This approach received increased attention over the last decade, when more scholars started testing language-related perceptual abilities of non-human animals. These studies have led to new insights on the abilities of animals to, for example, discriminate between phonological items and to generalize simple

rules. There has been no convincing evidence of animals learning abstract relations between arbitrary sounds (reviewed by 42). Further studies on how animals are perceiving sounds and abstract relations may shed a light on the evolution of language.

The work presented in this thesis is part of a collaborative project between developmental linguists, computational linguists and behavioural biologists. Within this project, we compare the cognitive mechanisms involved in language perception in infants, adults and two bird species. This allows us to explore how humans learn a language, which cognitive abilities are crucial and whether these abilities are shared with other species. Furthermore, we develop computational models for artificial grammar learning to untangle how different abilities influence language learning.

To determine which cognitive abilities are most crucial to learn a language, one can observe human infants in their language learning phase. Even though infants do not yet speak, their first year is crucial for their ability to understand and produce language later in life. It is in this year that they learn the structure of their native language. They learn the meaning of words and start recognizing the rhythmicity of speech (43,44). Furthermore, in this first year infants learn the prosody of sentences and the order and segmentation of words (45-47). The cognitive perceptual abilities shown by young infants are a good starting point for comparative research. These abilities are crucial for learning and understanding language and it might also be these abilities that were at the basis of the evolution of language.

In the next sections, I will describe five cognitive mechanisms related to language learning that were studied in the work presented in this thesis. There are two main reasons why we studied these five mechanisms. First, an early onset of cognitive abilities in humans could indicate that these traits are crucial for language. These traits might have an evolutionary history that predates language, which might be indicated by their presence in other species. And second, because of more pragmatic reasons. Our main study species, the zebra finch, is known to be able to memorize specific positions of sounds, transitions between elements and phonological features of speech sounds (22,48,49). We thus designed experiments that continued to explore these perceptual abilities further.

The first sections describe more basic abilities of language perception. How are sounds in a string memorized? Each sound can be memorized based on their ordinal position, or can be grouped together with some of the sounds heard earlier and later. We continue with an exploration of perceptual grouping biases that might be present

when sounds in long strings are alternating in prosodic stress. Next, we explore the perception of these stress or prosodic patterns in speech. We then continue with the more complex abilities related to abstract pattern learning. When perceiving strings of sounds, can they perceive the patterns that are underlying the strings and generalize these patterns to new sounds? We study this in a more musical setting, with the perception of different rhythms, and in a more linguistic setting, with the perception of different grammar rules.

String segmentation

To be able to determine the structural organization of a sentence, infants have to know which sound units form one word. Spoken language does not contain the reliable spaces between words as written language does. Therefore, infants have to rely on several other cues to segment a long sound sting into meaningful segments. In a pioneering study, Saffran and colleagues found that infants can trace the statistical probabilities between sound items and use this to segment strings (28). For example, in the utterance "*green apple*" the probability of "*green"* being followed by *"ap"* is rather low. You also hear "*green grass"* or "*red apple*". The probability that "*ap"* is followed by "*ple*" is much higher, as they always occur together, making it more probable that "*apple"* is a word and not "*greenap"*. To test this systematically, infants were exposed to a string of nonsense syllables. Words in this string were created by having a statistical probability of 1 between two syllables of a word. The probability between the last syllable of one word and the first one of the next was lower, as the words appeared in different orders. However, the silence between two syllables was always of identical length, whether this was within a word or between words. A string could be "*ABCDEFGHIABCGHIDEF…"*, where the probability of *A* being followed by *B* is 1, but the probability of *C* being followed by *D* is only 0.5. After two minutes of exposure to this syllable string, infants preferred the high probability words over the low probability words. Indicating that they were able to use the statistical probabilities between syllables to segment the string into words. Further investigation showed that, if present, infants can also use the pauses between syllables and the prosodic stress pattern of words to determine word boundaries (50-52). Although crucial for language learning, this ability did not specifically evolve for language perception. It is also present in the visual domain, where infants can segment strings of different images into logical units (52,53). String segmentation has thus far rarely been studied in the animal kingdom. Both cotton-top tamarins and rats seem to be able to use the co-occurrence of syllables in a string to segment strings (54,55). Zebra finches can detect the co-occurrences between elements, but it remains to be explored whether they can use this to segments strings.

Perceptual grouping

Infants can use the prosodic stress pattern of words to segment long sound string, which implies that they have a representation of the natural word stress of their language and perceptually organize sounds into groups that follow this stress pattern (28,56). Depending on the language, most words are either pronounced with initial stress (trochees) or with final stress (iambs). When human adults or infants hear sound items that are alternating in pitch or amplitude, they group these as trochees; a stress sound is followed by an unstressed sound. However, when the sounds are alternating in their duration, humans group these either as trochees or iambs depending on their native language (57-60). This strong tendency to perceive these rhythmic groups is called the "iambic/trochaic law" and aids the infants in segmenting speech strings into words (61,62). The perceptual grouping bias is not only present in the acoustic domain. When observing different visual items, humans also tend to group the items into pairs with either initial or final prominence (63). It was hypothesized that the iambic/trochaic grouping bias might be specific to humans. However, this may not be the case, as one study found that rats also group tones alternating in pitch as trochees (64). Rats showed no grouping bias towards tones alternating in duration. This study suggests that other species may also have perceptual grouping biases comparable to humans, indicating that this may have been a perceptual primitive for the evolution of speech and language.

Prosody perception

Prosody is paralinguistic information, created by stressing certain syllables by an increased frequency (pitch), amplitude or duration. Prosody can be used to determine which units form words, but changes in prosody can also alter the meaning of a sentence or reveal the emotional state of the speaker. These aspects of speech are very salient to humans. Newborns and older infants can already discriminate between languages based on only the prosodic pattern (44,65-68). This effect can also be found in their production: French and German newborns cry in the prosodic pattern of their own language (rising in French, falling in German) (69). Several non-human animals are also able to discriminate between the prosodic patterns of sentences (68,70,71). In different discrimination tasks, they manage to either discriminate between different languages or between different emotional prosodic patterns. The question remains, however, which prosodic features are used to make these discriminations.

Rhythm perception

Humans detect rhythm in music and can move their body to the beat of a song. Beat perception is shown in all age classes and also in people with little musical experience. Newborns already respond to the beat in a string of sounds and show increased brain activity when an expected the beat is absent (72). Furthermore, 5–7 month old infants can learn to discriminate between two short rhythmic sequences and generalize this rhythm to new sounds (73). However, infants learn to recognize a new rhythmic pattern faster than adults do (74), which indicates that there is a sensitive period for rhythm perception. Rhythm perception is not limited to the musical domain; different languages also have different rhythmic patterns in speech production. Not surprisingly, infants readily discriminate between the rhythmic patterns of different languages, which might assist in language learning (68,75).

Perceiving a particular rhythm or regularity, requires the participant to form an abstract representation of the rhythmic pattern. Only if that representation is formed, is the listener able to generalize the rhythm to new sounds or to faster or slower examples of the same rhythm. This means rhythm perception and generalization requires a similar ability to learn the abstract pattern of sound strings, just like in artificial grammar learning.

The study of rhythm perception in non-human animals was boosted when a sulphur-crested cockatoo was discovered that could entrain to different rhythms and was able to adjust its movements when the song was played faster or slower (76,77). This discovery led to the hypothesis that rhythm perception might be linked to vocal learning (8,9,76,77). Schachner and colleagues (76) conducted a survey of YouTube video's that seemed to strengthen the hypothesis that it were the vocal learners that could entrain to rhythms. More recently, a Californian sea lion and a bonobo showed rhythmic movements related to a regular sound pulse (78,79). These animals are not known to be vocal learners, suggesting that there might be more to rhythm perception than just vocal learning.

Artificial grammar learning

One of the pillars of language learning is the ability to learn grammatical rules that determine the correct sequence of words. Crucial to grammar learning is the ability to generalize abstract rules to novel utterances. To test this ability in a controllable experiment, Reber designed the artificial grammar learning (AGL) paradigm (31). Ever since, the AGL paradigm has been employed in many studies with human adults and infants in both the acoustic and visual domain (30,80-82). In experiments with human

infants, participants were familiarized or habituated with a set of artificially created sound strings that all follow the same grammatical rule. These grammar rules can be simple, like XYX or XXY (e.g. 25), they can have non-adjacent dependencies (e.g. 83,84), or they can be an artificial language constructed to follow more complex grammatical rules (e.g. 30). After a familiarization period, infants are presented with grammatical and ungrammatical structures. If the infants learned the grammar, they should discriminate between grammatical and ungrammatical structures, even if these structures consist of novel items. The infants show this by, for instance, a difference in looking times towards the two structure types. As an example, in a famous study by Marcus and colleagues (25) infants were familiarized with triplets containing either two successive syllables that were the same (XXY) or identical syllables on the two edges of the triplet (XYX). After two minutes of familiarization with one grammar, eight month old infants displayed an increased interest in the novel grammar, indicated by a longer looking time when sounds were played in an unfamiliar grammar compared to the same sounds in the familiarized grammar. This suggests that the infants formed an abstract representation of the familiarization grammar and compared the novel triplets to this template. Both the domain generality and the early onset of this abstraction ability indicate that it might have preceded language evolution. A hypothesis that would be strengthened if a non-human animal could learn similar abstract grammars.

All mechanisms mentioned above are involved in infant language learning, but have only been marginally studied in non-human animals. In order to gain insight into the perceptual abilities that might have been at the basis of language evolution, we need to know to whether and to which extend these abilities are present in non-human animals.

Which species to study?

When studying the shared principles underlying language perception, it might seem like a logical choice to study animals to which we are most closely related, great apes. These animals share many complex cognitive abilities with humans. For example, they can plan for the future (85,86) and understand causation (87). However, primates are not known to be vocal learners. This means that their vocalizations are not learned and, in general, less variable and elaborate than those of vocal-learning species. The cognitive abilities required for language and music perception are based on the ability to perceive and recognize different sounds. Vocal learners have to pay close attention to the sounds that they are copying, indicating that their auditory perception is well-developed. Furthermore, the subjects need to form abstract representations of structures and rules that are provided in sequences of various sound units. Whether it

is detecting a rhythm, a prosodic pattern or a syntactic structure, the observer has to find the abstract regularity that underlies the stimuli and generalize this to novel sounds.

There are animal species that need to pay attention to the regularities in conspecific vocalizations. Birds produce a variety of complex vocalizations, but are able to correctly discriminate between individuals of their own species and other species. For example, a nightingale can produce up to 200 different songs (88). Nevertheless, other nightingales recognize another nightingale as belonging to the same species. This means that these animals must be able to generalize a particular feature of the vocalizations to new songs. The same goes for animals with less variable vocalizations. Zebra finches produce one stereotyped song per individual (89,90). When a bird encounters a new individual with a song it never heard before, it can still recognize it as belonging to the same species.

Darwin already noted the parallels between human language and birdsong: "*The sounds uttered by birds offer in several respects the nearest analogy to language*" (91). These parallels have been studied in more detail over the last decades (e.g. 92). Most bird species learn their vocalizations early in life and require vocal input in this period, just like humans. Furthermore, both humans and vocal learning birds have a sensorimotor phase in which they already utter a premature version of their species specific vocalizations. In humans we call this the "babbling phase" and in birds it is referred to as "subsong" or "plastic song". Furthermore, the neural substrates related to song learning in vocal learning birds show analogies to the human system related to language learning (93,94). Hence, the ability to perceive a wide range of vocalizations as belonging to one species, the developmental and neural ability for vocal learning, and the current knowledge on the sounds they can perceive and discriminate between, makes birds an excellent group for comparative studies on language perception.

Zebra finches are the most well-studied songbird species. Both males and females produce social calls (short, mostly non-learned vocalizations), but only the males produce songs (long, variable, learned vocalizations). Each male learns one particular song early in life, a mixture of elements from his fathers' song and elements from other sources (95). These songs hardly change over the lifetime of an individual, making it clear markers of individual differences. The songs of zebra finches reveal information about their identity and females use the songs in their partner preference (96,97). The perceptual sensitivity of the zebra finches shows that females can use song structure to recognize an individual, and can use song rate and song complexity to gain more general information from the song, like indicators of the males' fitness.

Besides songbirds, parrots are also vocal learners. In contrast to the zebra finches, parrots are open-ended learners, meaning that they continue to learn new songs throughout their lives (98). One of the smaller parrots are the budgerigars. These animals have been extensively studied for their perception of human speech sounds (e.g. 99-101). They have similar sensitivities to humans to detect phoneme differences (101,102). This shows that besides being able to recognize an individual with high variability in the song, they can recognize the different sounds of human speech.

The above described vocal and perceptual abilities of zebra finches and budgerigars make them excellent candidates for comparative studies on language and music perception. They are both well studied and can perceive and discriminate different types of sounds, including human speech. It is of interest to compare the language perception abilities of these two species. For one, because they have different vocal learning abilities (open-ended or closed-ended). This influences the variation in their vocal communication, which might give rise to different perceptual abilities. Furthermore, although not specifically tested, zebra finches and budgerigars might have different cognitive abilities as they belong to two different clades (songbirds and parrots). By comparing two, very different, vocal learning avian species, we can gain more insight into the traits that might predict abstract pattern learning.

Thesis focus

In this thesis I focus on the question of which aspects of language perception are shared between humans and non-human animals. More specifically, I test two vocal-learning bird species, zebra finches and budgerigars, on their abilities to abstract and generalize acoustic structural patterns.

Thesis outline

This thesis describes a set of experiments on various perceptual abilities of zebra finches and budgerigars that are relevant to detect patterns in music and language. The experiments concentrated on perceptual abilities that are present during early language development in infants. Furthermore, the experiments all had a level of sound complexity that has been shown to be perceivable to the birds (as discussed earlier in this introduction). We tested whether zebra finches and budgerigars have the cognitive abilities required to segment sound strings based on transitions, perceptually group stress alternations, generalize prosodic cues, abstract rhythmicity and process grammar rules.

CHAPTER 2 describes a study on the string segmentation strategies of zebra finches. When infants learn their native language, they have to quickly learn which segments of a string of sounds (a sentence) form a word. As natural speech does not contain systematic pauses between words, this task is rather complex. Infants use both the co-occurrence of the sound items as well as pauses between words and other prosodic cues to segment a sound string into logical segments. We tested how zebra finches segment long strings of zebra finch song elements. One group was trained with strings of triplets with equally short pauses between all elements in the string. A second group was trained on strings with elongated pauses between triplets and short pauses between elements within a triplet. Thus, they both could use the co-occurrence of elements within a triplet, but additionally the second group could also use the information provided by the pauses. After this training, the zebra finches were retrained on full triplets as they occurred in the string, combinations of adjacent triplets in the string and random combinations of elements. If they segmented the training strings, they are expected to recognize the full triplets, but not the combinations, as belonging to one of the two training strings.

CHAPTER 3 describes experiments that tested whether zebra finches share the perceptual grouping biases for prosodic variation that are common in humans. When humans hear a string of tones that alternate in a prosodic feature like pitch, duration or amplitude, they have the natural tendency to perceptually group the tones in duplets with initial (trochaic) or final (iambic) stress. The perceptual bias to group alternations in pitch and amplitude as trochees is universal across languages and is also found in rats. Grouping duration alternations is dependent on the participants' native language and has not been found in other animals. In this study, zebra finches were trained to give a different behavioural response to trochees than to iambs. One group of zebra finches heard these stress patterns with pitch modulations, the other group with duration modulations. After training they were exposed to ambiguous long strings of alternating tones. If the birds perceptually group these strings as a concatenation of iambs or trochees, they are expected to respond to them in the same way as to the trochees or iambs of the training.

CHAPTER 4 continues to study the perception of prosodic cues by zebra finches. Specifically, we study how zebra finches respond to stress patterns of natural human speech. Prosody in human speech is a modification of the pitch, duration and amplitude of a syllable. This paralinguistic information helps infants to segment long sound strings into words. The zebra finches in this study were trained to discriminate between quadruplets of human speech syllables with an XXYY and XYXY structure with prosodic stress on either the first or the last syllable. Thus, these training

quadruplets contained both a structural as well as a prosodic difference. Subsequent tests were conducted to see whether the zebra finches responded more to the prosody or to the structure of the quadruplets, whether they could generalize the prosodic pattern to new syllables and which prosodic cues the birds were using when discriminating.

The study in CHAPTER 5 explores whether zebra finches and budgerigars can recognize and generalize the rhythmic pattern of tonal strings. One of the main hypotheses on rhythm perception is that vocal learning is required. Nevertheless, this hypothesis is mainly based on the many parrot species that show rhythmic entrainment and regularity perception. By comparing a songbird, the zebra finch, with a parrot species, the budgerigar, we aim to determine whether parrots are especially good at perceiving rhythms, or whether this ability is also present in other species. The birds were trained to discriminate between a regular and an irregular beat pattern. Subsequently, they were tested with slower and faster versions of these beat patterns.

CHAPTER 6 is an opinion piece in which we relate the rhythm perception abilities of zebra finches to two recently published papers on rhythms in their vocal production. The study of chapter 5 and other studies have shown that zebra finches can discriminate between strings with regular and irregular intervals. Recently, analyses of the songs and contact calls of zebra finches showed that these vocalizations often have a fixed rhythmic pattern. Furthermore, specific nuclei of the zebra finch song system are involved in both producing and detecting rhythmicity. This relation of rhythmicity in perception and production shows that zebra finches might be the ideal model species for the study on rhythm.

CHAPTER 7 describes experiments in which both zebra finches and budgerigars are tested for their artificial grammar learning abilities. Grammar learning is one of the fundamental aspects of language perception, as it allows the listener to correctly determine the meaning of a sentence. Seven month old infants are able to learn the abstract relationship between syllables organized in XYX or XXY grammatical structures, where the X is one syllable type and the Y another syllable. They can generalize these structural patterns to triplets that consist of new syllables, showing that they indeed learned the abstract grammatical structure. To test the uniqueness of this ability, we trained zebra finches and budgerigars to discriminate between XYX and XXY structures. Subsequently they were tested with triplets with the same structures but consisting of combinations of previously-heard sounds or completely new sounds.

CHAPTER 8 summarizes the work presented in this thesis and discusses the implications of these results.

REFERENCES

1. Chomsky, N. (1995). The minimalist program. MIT Press, Cambridge, MA.

2. Chomsky, N. (2005). On Phases. MIT Press, Cambridge, MA.

3. McNeill, D. (1992). Hand and mind. Chicago: University of Chicago Press.

4. Pinker, S. & Bloom, P. (1990). Natural language and natural selection. Behavioral and brain sciences, 13(4), 707-727.

5. Kirby, S. & Hurford, J.R. (1997). Learning, culture and evolution in the origin of linguistic constraints. In ECAL97 (Husbands, P. and Harvey, I., eds) pp. 493–502, MIT Press.

6. Jackendoff, R. (1997). The architecture of the language faculty. MIT Press.

7. Tyack, P. L. & Sayigh, L. S. (1997). Vocal learning in cetaceans. In: Social influences on vocal development, eds., C.T. Snowdon & M. Hausberger. Cambridge University Press.

8. Patel, A. D., Iversen, J. R. & Bregman, M. R. (2009a). Experimental evidence for synchornizatino to a musical beat in a nonhuman animal. Curr. Biol. 19, 827–830.

9. Patel, A. D., Iversen, J. R., Bregman, M. R. & Schulz, I. (2009b). Studying synchronization to a musical beat in nonhuman animals. Ann. N. Y. Acad. Sci. 1169, 459–469.

10. Reichmuth, C. & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. Current opinion in Neurobiology, 28, 66-71.

11. Boughman, J.W. (1998). Vocal learning in greater spear-nosed bats. Proc. R. Soc. Lond. B. 265, 227-233.

12. Poole, J.H., Tyack, P.L., Stoeger-Horwath, A.S. & Watwood, S. (2005). Animal behaviour: Elephants are capable of vocal learning. Nature, 434, 455-456.

13. Baptista, L.F. & Schuchmann, K.L. (1990). Song learning in the Anna hummingbird (Calypte anna). Ethology 84, 15-26.

14. Pepperberg, I. M. (1994). Vocal learning in gray parrots (Psittacus erithacus): Effects of social interaction, reference, and context. Auk 111, 300–313.

15. Marler, P. & Peters, S. (1977). Selective vocal learning in a sparrow. Science, 198, 519-521.

16. Hauser, M.D., Chomsky, N. & Fitch, W.T. (2002). The Faculty of Language: What Is It, Who Has It, and how did it evolve? Science, 298, 1569.

17. Fitch, W. T., Hauser, M. D. & Chomsky, N. (2005). The evolution of the language faculty: clarifications and implications. Cognition, 97, 179–210.

18. Miller, J. L. & Eimas, P. D. (1983). Studies on the categorization of speech by infants. Cognition, 13(2), 135–165.

19. Grieser, D. & Kuhl, P.K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. Developmental Psychology, 25, 577-588.

20. Lieberman, P. (1975). On the Origins of Language. An Introduction to the Evolution of Human Speech. Macmillan.

21. Liberman, A.M. (1982). On finding that speech is special. Am Psychol, 37:148-167.

22. Ohms, V.R., Gill, A., van Heijningen, C.A.A., Beckers, G.J.L. & ten Cate, C. (2010). Zebra finches exhibit speaker independent phonetic perception of human speech. Proc. R. Soc. B 277, 1003– 1009.

23. Ohms, V.R., Escudero, P., Lammers, K. & ten Cate, C. (2012). Zebra finches and Dutch adults exhibit the same cue weighting bias in vowel perception. Anim. Cogn. 15, 155–161.

24. Kriengwatana, B., Escudero, P., Kerkhoven, A. H. & ten Cate, C. (2015). A general auditory bias for handling speaker variability in speech? Evidence in humans and songbirds. Frontiers in Psychology, 6, 1234.

25. Marcus, G.F., Vijayan, S., Rao, S.B. & Vishton, P.M. (1999). Rule learning by seven-month-old infants. Science 283(5398):77-80.

26. Chambers, K.E., Onishi, K.H. & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. Cognition, 87(2), B69–B77.

27. Monaghan, P. & Christiansen, M. H. (2010). Words in puddles of sound: Modelling psycholinguistic effects in speech segmentation. Journal of Child Language 37:545–64.

28. Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-monthold infants. Science 274:1926–28.

29. Rowland, C., Chang, F., Ambridge, B., Pine, J.M. & Lieven, E.V. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. Cognition 125(1); 49–63.

30. Gomez, R.L. & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70:109–35.

31. Reber, A.S. (1967). Implicit Learning of Artificial Grammars. J Verb Learn Verb Behav, 6, 855-863.

32. Gomez, R. L. & Gerken, L. (2000). Infant artificial language learning and language acquisition. Trends Cogn. Sci. 4, 178–186.

33. Standley, J.M. & Madsen, C.K. (1990). Comparison of infant preferences and responses to auditory stimuli: Music, mother, and other female voice. Journal of Music Therapy, 27(2), 54-97.

34. Soley, G. & Hannon, E.E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. Developmental psychology, 46(1), 286-292.

35. Miller, G.F. (2000). Evolution of human music through sexual selection. In N.L. Wallin, B. Merker & S. Brown (eds), The origins of music (pp. 329-360). MIT Press, Cambridge, MA.

36. Cross, I. (2007). Music and cognitive evolution. In R.I.M. Dunbar & L. Barrett (Eds.), Oxford handbook of evolutionary psychology (pp.649–667). Oxford, England: Oxford University Press.

37. Merker, B. (2000). Synchronous chorusing and human origins. In N.L. Wallin, B. Merker, & S. Brown (eds.), The origins of music (pp.315–327). MIT Press, Cambridge, MA.

38. Honing, H. & Ploeger A. (2012). Cognition and the evolution of music: pitfalls and prospects. Topics Cogn. Sci, 4(4), 513-524.

39. McDermott J. (2008). The evolution of music. Nature 453:287–8.

40. Fitch, W.T. (2006). The biology and evolution of music: a comparative perspective. Cognition 100: 173. Miller, G. F. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, & S. Brown (Eds.), The origins of music (pp. 329–360). Cambridge, MA: MIT Press.

41. Pinker, S. (1999). How the mind works. Annals of the New York Academy of Sciences, 882, 119-127.

42. ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367, 1984–1994.

43. Markman, E.M. (1994). Constraints on word meaning in early language acquisition. Lingua, 92, 199-227.

44. Nazzi, T., Jusczyk, P.W. & Johnson, E.K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. Journal of memory and language, 43(1), 1-19.

45. Christophe, A., Mehler, J. & Sebastian-Galles, N. (2001). Perception of prosodic boundary correlates by newborn infants. Infancy, 2(3), 385-394.

46. Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of memory and language, 44(4), 548-567.

47. Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. Journal of memory and language, 57(1), 24-48.

48. Chen, J. & ten Cate, C. (2015). Zebra finches can use positional and transitional cues to distinguish vocal element strings. Behav Process. 117, 29-34.

49. van Heijningen, C.A., Chen, J., van Laatum, I., van der Hulst, B. & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? Anim Cogn, 16:165-175.

50. Morgan, J.L. (1996). A rhythmic bias in preverbal speech segmentation. Journal of Memory and Language, 35, 666–688.

51. Thiessen, E.D. & Saffran, J.R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. Developmental Psychology, 39, 706–716.

52. Saffran, J.R., Johnson, E.K., Aslin, R.N. & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. Cognition, 70, 27–52.

53. Turk-Browne, N.B., Jungé, J.A. & Scholl, B.J. (2005). The automaticity of visual statistical learning. J. Exp. Psychol. Gen. 134, 552–564.

54. Hauser, M.D., Newport, E.L. & Aslin, R.N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. Cognition, 78, B53–B64.

55. Toro, J.M. & Trobalon, J. (2005). Statistical computations over a speech stream in a rodent. Perception & Psychophysics 67:867.

56. Goodsitt, J.V., Morgan, J.L. & Kuhl, P.K., (1993). Perceptual strategies in prelingual speech segmentation. Journal of Child Language 20, 229–252.

57. Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. In Proceedings of the Thirteenth meeting of the Berkeley Linguistics Society, ed. by M. Niepokuj, et. al., 429–446. Berkeley, CA: Berkeley Linguistics Society.

58. Iversen, J.R., Patel, A.D. & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. The journal of the Acoustical Society of America 124(4), 2263-2271.

59. Yoshida, K.A., Iversen, J.R., Patel, A.D., Mazuka, R., Nito, H., Gervain, J., et al. (2010). The development of perceptual grouping biases in infancy: A Japanese–English cross-linguistic study. Cognition, 115, 356–361.

60. Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T. & Hohle, B. (2013). Native language affects rhythmic grouping of speech. The Journal of the Acoustical Society of America, 134(5), 3828–3843.

61. Bion, R.A., Benavides-Varela, S. & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. Language and Speech, 54(1), 123–140.

62. Hay, J.S.F. & Saffran, J.R. (2012). Rhythmic grouping biases constrain infant statistical learning. Infancy, 17(6), 610–641.

63. Peña, M., Bion, R.A.H. & Nespor, M. (2011). How modality specific is the iambic-trochaic law? Evidence from vision. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37, 1199–1208.

64. de la Mora, D., Nespor, M. & Toro, J.M. (2013). Do humans and non-human animals share the grouping principles of the Iambic – Trochaic Law? Att Percept Psychophys, 75, 92-100.

65. Jusczyk, P.W., Cutler, A. & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. Child Development, 64(3), 675–687.

66. Mehler, J. & Christophe, A. (1995). Maturation and learning of language during the first year of life. In M. S. Gazzaniga (Ed.), The Cognitive Neurosciences (pp. 943-954): Bradford Books / MIT Press.

67. Nazzi, T., Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. Journal of Experimental Psychology: Human Perception and Performance, 24(3), 756-766.

68. Ramus, F., Hauser, M.D., Miller, C., Morris, D. & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. Science, 288, 349– 351.

69. Mampe, B., Friederici, A. D., Christophe, A. & Wermke, K. (2009) Newborns' cry melody is shaped by their native language. Current Biology 19:1994–97.

70. Naoi, N., Watanabe, S., Maekawa, K. & Hibiya, J. (2012). Prosody discrimination by songbirds (Padda oryzivora). PLoS ONE 7, e47446.

71. Toro, J.M., Trobalon, J.B. & Sebastian-Galles, N. (2003). The use of prosodic cues in language discrimination tasks by rats. Anim. Cogn. 6, 131–136.

72. Winkler, I., Haden, G., Ladinig, O., Sziller, I. & Honing, H. (2009). Newborn infants detect the beat in music. Proc. Natl. Acad. Sci. USA, 106, 2468–2471.

73. Trehub, S. E. & Thorpe, L. A. (1989). Infants' perception of rhythm. Categorization of auditory sequences by temporal structure. Canadian Journal of Psychology, 43, 217–229.

74. Hannon, E.E. & Trehub, S.E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. PNAS 102(35), 2639–12643.

75. Nazzi, T. & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. Speech Communication, 41, 233–243.

76. Schachner, A., Brady, T.F., Pepperberg, I.M. & Hauser, M.D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. Curr. Biol. 19, this issue, 831–836.

77. Patel, A.D. (2006). Musical rhythm, linguistic rhythm, and human evolution. Music Percept. 24, 99–104.

78. Cook, P., Rouse, A., Wilson, M. & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. J. Comp. Psychol. 127, 412–427.

79. Large E. W., Gray P. (2015). Spontaneous tempo and entrainment in a bonobo (pan paniscus). J. Comp. Psychol. 129(4), 317-328.

80. Christiansen, M.H., Conway, C.M. & Curtin, S. (2000). A connectionist single mechanism account of rule-like behavior in infancy. In: Proceedings of the 22nd Annual Conference of the Cognitive Science Society. Erlbaum.

81. Saffran, J.R., Pollak, S.D., Seibel, R.L. & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. Cognition 105:669–680.

82. Petersson, K. M., Folia, V. & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. Brain Lang. 120, 83–95.

83. Santelmann, L.M. & Jusczyk, P.W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. Cognition, 69, 105–134.

84. Gomez, R.L. & Maye, J. (2005) The developmental trajectory of nonadjacent dependency learning. Infancy 7, 183–206.

85. Mulcahy, N.J. & Call, J. (2006). Apes save tools for future use. Science 312, 1038–1040.

86. Osvath, M. & Osvath, H. (2008). Chimpanzee (Pan troglodytes) and orangutan (Pongo abelii) forethought: self-control and pre-experience in the face of future tool use. Anim. Cogn., 11, 661–674.

87. Visalberghi, E. & Tomasello, M. (1998). Primate causal understanding in the physical and psychological domains. Behavioural Processes 42(2–3):189–203.

88. Hultsch, H. & Todt, D. (1989). Song acquisition and acquisition constraints in the nightingale, Luscinia megarhynchos. Naturwissenschaften, 76(2), 83-85.

89. Immelmann K. (1969). Song development in the zebra finch and other estrildid finches. In: Hinde RA, editor. Bird Adult Song Plasticity Requires LMAN 27 vocalizations. Cambridge: Cambridge University Press. p 61–74.

90. Price P. (1979). Developmental determinants of structure in zebra finch song. J Comp Physiol Psychol 93:260–277.

91. Darwin, C. (1882). The Descent of Man and Selection in Relation to Sex. (Murray, London).

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

93. Jarvis, E.D. (2007). Neural systems for vocal learning in birds and humans: a synopsis. Journal of ornithology, 148, 35-44.

94. Bolhuis, J.J., Okanoya, K. & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. Nature Reviews Neuroscience, 11, 747-759.

95. Eales, L.A. (1985). Song learning in zebra finches: some effects of song model availability on what is learnt and when. Animal Behaviour, 33(4), 1293-1300.

96. Collins, S.A., Hubbard, C. & Houtman, A.M. (1994). Female mate choice in the zebra finch—the effect of male beak colour and male song. Behavioral Ecology and Sociobiology, 35(1), 21-25.

97. Holveck, M.J. & Riebel, K. (2007). Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three test contexts. Animal Behaviour, 74(2), 297-309.

98. Farabaugh, S.M., Linzenbold, A. & Dooling, R.J. (1994). Vocal Plasticity in Budgerigars (Melopsittacus undulates}: Evidence for Social Factors in the Learning of Contact Calls. Journal of Comparative Psychology, 108(1), 81-92.

99. Dooling, R.J., Okanoya, K. & Brown, S.D. (1989). Speech perception by budgerigars (Melopsittacus undulatus): The voiced-voiceless distinction. Perception & psychophysics, 46(1), 65-71.

100. Dooling, R.J. & Brown, S.D. (1990). Speech perception by budgerigars (Melopsittacus undulates): spoken vowels. Perception & psychophysics, 47(6), 568-574.

101. Dent, M.L., Brittan-Powell, Dooling, R.J. & Pierce, A. (1997). Perception of synthetic /ba/-/wa/ speech continuum by budgerigars (Melopsittacus undulates). J. Acoust. Soc. Am., 102, 1891.

102. Dooling, R.J., Best, C.T. & Brown, S.D. (1995). Discrimination of synthetic full−formant and sinewave/ra–la/continua by budgerigars (Melopsittacus undulatus) and zebra finches (Taeniopygia guttata). J. Acoust. Soc. Am. 97, 1839.

Michelle Spierings, Anouk de Weger & Carel ten Cate

Published in Animal Cognition, 2015, 18(4), 867-874.



2

PAUSES ENHANCE CHUNK RECOGNITION IN SONG ELEMENT STRINGS BY ZEBRA FINCHES

ABSTRACT

When learning a language, it is crucial to know which syllables of a continuous sound string belong together as words. Human infants achieve this by attending to pauses between words or to the co-occurrence of syllables. It is not only humans that can segment a continuous string. Songbirds learning their song tend to copy 'chunks' from one or more tutors' songs and combine these into their own song. In the tutor songs, these chunks are often separated by pauses and a high co-occurrence of elements, suggesting that these features affect chunking and song learning. We examined experimentally whether the presence of pauses and element co-occurrence affect the ability of adult zebra finches to discriminate strings of song elements. Using a go/no-go design, two groups of birds were trained to discriminate between two strings. In one group (Pause group), pauses were inserted between co-occurring element triplets in the strings, and in the other group (No-pause group), both strings were continuous. After making a correct discrimination, an individual proceeded to a reversal training using string segments. Segments were element triplets consistent in co-occurrence, triplets that were partly consistent in composition and triplets consisting of elements that did not co-occur in the strings. The Pause-group was faster in discriminating between the two strings. This group also responded differently to consistent triplets in the reversal training, compared to inconsistent triplets. The No-pause group did not differentiate among the triplet types. These results indicate that pauses in strings of song elements aid song discrimination and memorization of co-occurring element groups.

INTRODUCTION

Learning which syllables of a continuous speech stream belong together as words is one of the first challenges that human infants face when they are acquiring a language. In order to do so, infants attend to the pauses between words (1-4). However, pauses are not always reliable and can occur both between and within words. Another way of detecting regularities is by paying attention to the transitions between syllables (5-7). Syllables that occur together more often and have a higher transitional probability are more likely to form a word. Infants also use this feature to correctly segment speech streams. Computational models support the hypothesis that transitional information can be sufficient for correct word segmentation (8, 9).

Humans are not the only animals that segment longer acoustic sequences into smaller units. Other vocal learners, like songbirds, copy groups of song elements or song types from tutor songs and combine these in their own song. Nightingales, for instance, learn 'packages' of a few song types and combine these as units in their own song sequences (10). When exposed to many song types, young nightingales tend to copy groups of song types that often occur together or that are surrounded by longer pauses (11, 12). This shows that the packages they learn are based on both proximity of song types in the tutors' song, as well as on pauses between song types.

Songbirds with less vocal variation also show a tendency to copy chunks from their tutors' song. Zebra finches often copy groups of elements instead of single elements and can combine chunks from different tutors into their own song (13, 14). In the tutor songs, these chunks are separated by relatively long silent intervals (14). Interrupted songs are terminated most often at the end of chunks and respiratory patterns show inhales and exhales at chunk edges (15, 16). Bengalese finches also seem to perceive songs as a composition of chunks (17) and combine chunks in their own song (18). The elements within these chunks co-occur more often and have shorter pauses between them compared to elements of adjacent chunks (18-20). These studies imply that pauses between groups of elements and co-occurrence of elements within a group affect the memorization of songs and song segments in young birds. However, in natural songs, pauses and element co-occurrence often coincide, making it hard to establish the importance of each factor. Also, it is unknown whether pauses and element co-occurrence affect song memorization in adult birds.

In the current study, we examine the role of both pauses and co-occurrence on song memorization in adult zebra finches. To this end, we trained the birds to discriminate artificially edited strings of song elements. Zebra finches are able to pay attention to both position and co-occurrence when learning sequences of elements (21) and can identify short strings of identical song elements based on differences in element sequence (22-25). In the current experiment, elements in the training strings are arranged in triplets based on co-occurrence, or -in the second experimental group- based on pauses between the triplets as well as element co-occurrence.

METHODS

Subjects

Twenty-eight Zebra finches (14 males, 14 females; ages 175–280 days post hatching) were used for this study. All birds were bred and reared at Leiden University and had not been used in experiments before. Half of the birds were assigned to the Pause-group, the other half to the No-pause group (seven males and seven females in both groups; age Pause-group: M = 217, SD = 30, age No-pause group: M = 215, SD = 34). Before the experiment, the zebra finches were housed in single sex groups on a 13.5 L:10.5 D schedule at 20–22 °C. During the experiment, water, grit and cuttlebone were available ad libitum. Food was used as reinforcement and only available after a correct trial. The birds' food intake was monitored daily and additional food was given when necessary.

Operant cages

The experimental setup was identical to that used by Spierings and ten Cate (26). All experiments were conducted in an operant conditioning cage [70 (l), 30 (d), 45 (h) cm]. Each cage was in a separate sound-attenuated chamber and illuminated by a fluorescent tube that emitted a daylight spectrum on a 13.5 L:10.5 D schedule. A speaker (Vifa 10BGS119/8) was located 1 m above the cage. The cage was made from wire mesh except for the floor and a plywood back wall which supported two pecking keys with LED lights. A food hatch was located in between these two keys, easily accessible to the birds. Pecking the left key (sensor 1) elicited a stimulus and illuminated the LED light of the key on the right (sensor 2). Depending on the sound, the bird had to peck sensor 2 or had to withhold its response. A correct response resulted in access to food for 10 s and an incorrect response led to 15 s of darkness.

Training

The experiment consisted of one shaping phase and two training phases. Shaping was required to familiarize the subject with the setup. During the first training phase (string discrimination training), the birds had to discriminate between two strings. The second training phase (segments reversal training) was a reversal training in which the birds were trained on several specific combinations of elements from the initial strings. During this reversal training, the corresponding feedback was reversed; combinations of elements that originated from the go-string were now no-go items and vice versa.

Shaping. Zebra finches were first trained in the go/no-go task without exposure to the experimental stimuli. The birds received a conspecific song as the go-stimulus and a pure tone as the no-go stimulus. Each day, the discrimination between the stimuli by each individual was calculated as a percentage correct score (%C) as follows: (correct go responses + correct no-go rejections)/total number of trials. If a bird made no mistakes in the discrimination by always responding to a go-stimulus and never to a no-go stimulus, their %C would be 1. For example, in 20 trials, this would be (10 go responses + 10 no-go rejections)/20 trials. Performance at random results in an %C of 0.5, for example, when a bird pecked to a go-stimulus in only 50 % of the cases and also to a no-go stimulus in 50 % of the cases over 20 trials: (5 go responses + 5 no-go rejections)/ 20 trials = 0.5. A fully incorrect discrimination (only responding to no-go stimuli) would lead to a %C of 0. This shaping phase lasted until the zebra finch reached the shaping criterion of %C >0.8 for three consecutive days, after which the training switched to the string discrimination training.

String discrimination training. All individuals were trained with one go and one no-go string of zebra finch song elements. For every block of 100 trials, the %C was calculated. A bird progressed to the segments reversal training after reaching the learning criterion of eight consecutive blocks of 100 trials with %C >0.8.

Segments reversal training. The segments reversal training consisted of triplets of elements which were reinforced with reversed contingencies compared to the string training. Segments from the go-string were now reinforced as being no-go stimuli and segments from the no-go string as go-stimuli. For example, a go-response to a triplet from the previous go-string would now result in 15 s of darkness, and a go-response to a triplet from the previous no-go string would result in 10 s of food access. This phase lasted for 3000 trials, independent of the birds' performance. The hypothesis underlying this reversal training was that the birds would have most difficulties with

reversing their response to triplets that had become associated strongly with the feedback during the string training. Overall, animals show an increase in incorrect responses after the contingencies of stimuli are reversed. However, the speed by which these new contingencies are learned might be influenced by several factors (for an overview, see 27). Therefore, we focus on the first 20 trials after the reversal only.

Stimuli

String stimuli. Two strings were constructed from 12 zebra finch song elements, originating from normal songs. All elements were chosen to be from different element categories and were equalized in amplitude, and the beginning and end of the elements were ramped (5 ms on each side) with Praat (28). Both strings consisted of four unique triplets, which were a fixed concatenation of three different song elements (for examples of a go and no-go string, see Fig. 1). The triplets of the go and no-go string shared the starting element, but the second and third element of the triplets were different in the two strings (Table 1). The strings were arranged with 20 ms pauses between adjacent elements. The No-pause experimental group was trained to discriminate between two of such strings. The Pause-group was trained using the same strings, but with prolonged pauses (80 ms) between the triplets (Fig. 1). The letters in Table 1 depict the 12 song elements that were used and their order in the string. To avoid pseudo replication, the element represented by each letter was different for each bird within a group. For instance, element 'A' could be a 'high trill' element for one bird, but would be a 'flat' element for another bird. Each string combination occurred both in the Pause-group and the No-pause group, resulting in fourteen different go and fourteen different no-go strings.

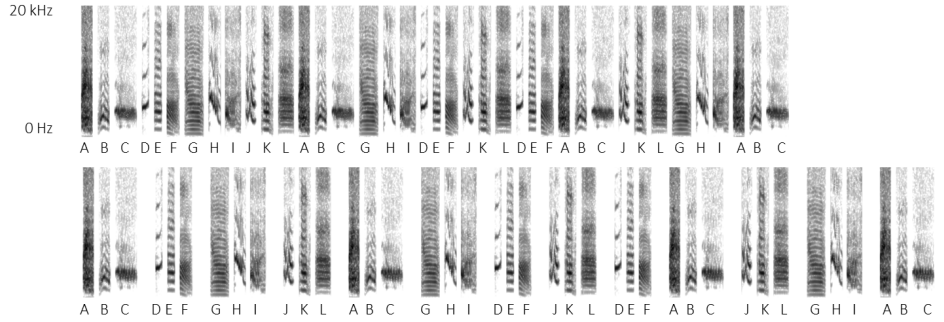


Fig. 1 Example of a string with and a string without pauses. Adjacent elements were separated by a 20 ms pause. In the Pause-condition a 80 ms pause separated the triplets.

Segment stimuli. Three different types of triplets were created for the reversal training, which we refer to as consistent triplets, partly consistent triplets and inconsistent triplets (Table 2). 1) Consistent triplets were a combination of three elements that had always occurred as a concatenated triplet in the training string. This means that these three elements had high co-occurrence and occurred in both experimental conditions without long pauses between them. 2) Partly consistent triplets were a combination of the last element of one triplet and the first two elements of the following triplet. This means that the co-occurrence of the first and the second element of a partly consistent triplet is lower than the co-occurrence of the second and third element. In the Pause-condition, there had also been a pause between the first and second element of a partly consistent triplet in the initial training strings. 3) Inconsistent triplets consisted of a combination of elements that had never occurred together during the string discrimination training. Therefore, there had not been co-occurrence between these elements in the training strings. For each triplet category (consistent triplets, partly consistent triplets and inconsistent triplets), two triplets were derived from the go-string and two from the no-go string of each bird (see Table 1 for a representation of the string stimuli, and Table 2 for a representation of the consistent triplets, partly consistent triplets and inconsistent triplets).

Table 1. The go and no-go strings as presented to the birds. Each string consisted of four different triplets, here shown with pauses between the triplets. The No-pause group received the elements in a continuous string. By organizing the triplets in this way, we created strong co-occurrence of song elements within a triplet and lower co-occurrence of elements of different triplets. The letters represent zebra finch song elements which were different for each individual in a group.

Go String
ABC DEF GHI JKL ABC GHI DEF JKL DEF ABC JKL GHI ABC
No-Go String
AFK DLB GIE JCH AFK GIE DLB JCH DLB AFK JCH GIE AFK

Table 2. The different triplets created for training. These triplets were different combinations of the elements from the string discrimination training.

	from Go String	**from No-Go String**
Triplets	JKL GHI	JCH GIE
Part-triplets	IJK LDE	EJC HDL
Non-triplets	ELH HJD	IHL LIG

Statistical analyses

String discrimination training. This training phase was completed when an individual reached a %C > 0.8 for eight successive blocks of 100 trials. The number of days and the number or trials needed to reach this criterion were measured for each individual. These measurements followed a normal distribution (number of trials after a log transformation), allowing us to analyze the results of the Pause-group and the No-pause group in a paired Student's t test. The groups were paired based on the similarity of the discrimination training strings. Every element combination was present once with pauses between the triplets and once without pauses.

Segments reversal training. For each individual, an average %C was calculated for the first 20 trials of each triplet category (consistent triplets, partly consistent triplets and inconsistent triplets). These scores were used to measure the first responses of the birds to these segments. The data were analyzed in a linear mixed effects model (LMNE) with %C as the dependent variable and condition (pauses or no pauses) and triplet category as independent variables. Individual was inserted as the random variable. Differences between the triplet categories were exposed with a post hoc Tukey test. Furthermore, we analyzed whether the %C of each triplet category per group deviated from random performance (%C = 0.5), with a one sided t test. The significance levels were corrected with a Bonferroni correction due to repeated analyses within one experimental group.

Correlations. In order to reveal possible correlations between the first and the second training phase, we ran a Pearson correlation test between the birds' %C of the three different triplet categories and the number of days and trials they needed to complete the first training phase.

RESULTS

String discrimination training

All zebra finches reached the discrimination criterion (3 days with %C > 0.8) within 11 days, with an average of 477 trials per day (SD = 109.6). The Pause-group achieved the discrimination earlier than the No-pause group (Pause: M = 5.14, SD = 1.61; No-pause: M = 7.07, SD = 2.07; t = -3.49, P = .004; Fig. 2). The Pause-group also needed fewer trials to reach the learning criterion (Pause: M = 2485, SD = 718; No-pause: M = 3171, SD = 1148; t = -2.23, P = .03).

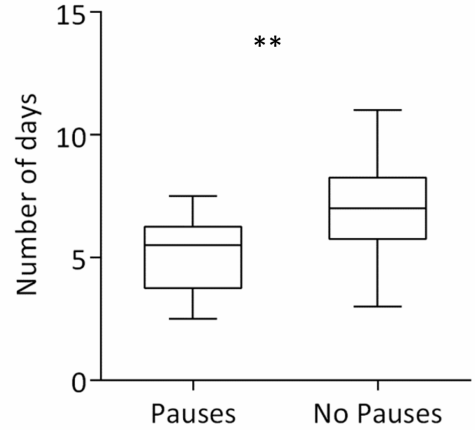


Fig. 2 Number of days needed to reach the learning criterion. Discrimination between strings was made more quickly with pauses between the triplets of the strings. The boxplots show the median (horizontal line) and first and third quartile of the data, with whiskers extending to the minimum and maximum values. Asterisks indicate a significant difference between the groups.

Segments reversal training

For each triplet category, the %C was calculated over the first 20 trials of the reversal training. The Pause-group responded differently to the consistent triplets (c-triplets) than to the partly consistent triplets (pc-triplets) and inconsistent triplets (ic-triplets) (lmm Tukey c-triplets vs. pc-triplets: $z = -3.01$, $P = .007$; c-triplets vs. ic-triplets: $z = -3.51$, $P = .001$; pc-triplets vs. ic-triplets: $z = -0.50$, $P = .87$; Fig. 3). The response to consistent triplets was significantly lower than random (=0.5), while the responses to both partly consistent triplets and inconsistent triplets did not deviate from random (c-triplets: $M = 0.46$, $SD = 0.06$, $P = .04$; pc-triplets: $M = 0.53$, $SD = 0.06$, $P = .17$; ic-triplets: $M = 0.54$, $SD = 0.08$, $P = .13$). The No-pause group showed no difference in response to consistent triplets, partly consistent triplets or inconsistent triplets (lmm Tukey c-triplets vs. pc-triplets: $z = -1.48$, $P = .30$; c-triplets vs. ic-triplets: $z = -0.74$, $P = .74$; pc-triplets vs. ic-triplets: $z = 0.74$, $P = .74$; Fig. 3). Neither of the groups showed an effect of sex (pauses: $P = .43$; no pauses: $P = .87$).

These results remained consistent over the first 100 trials of each triplet category. In the Pause-group, the responses to the consistent triplets stayed lower than the responses to the partly consistent and inconsistent triplets (lmm Tukey c-triplets vs. pc-triplets: $z = -2.70$, $P = .02$; c-triplets vs. ic-triplets: $z = -3.14$, $P = .005$; pc-triplets vs. ic-triplets: $z = -0.44$, $P = .90$; Fig. 4). The No-pause group continued to not show a difference in response to any of the triplet categories (lmm Tukey c-triplets vs. pc-triplets: $z = -1.16$, $P = .478$; c-triplets vs. ic-triplets: $z = -0.33$, $P = .94$; pc-triplets vs. ic-triplets: $z = 0.83$, $P = .68$; Fig. 4).

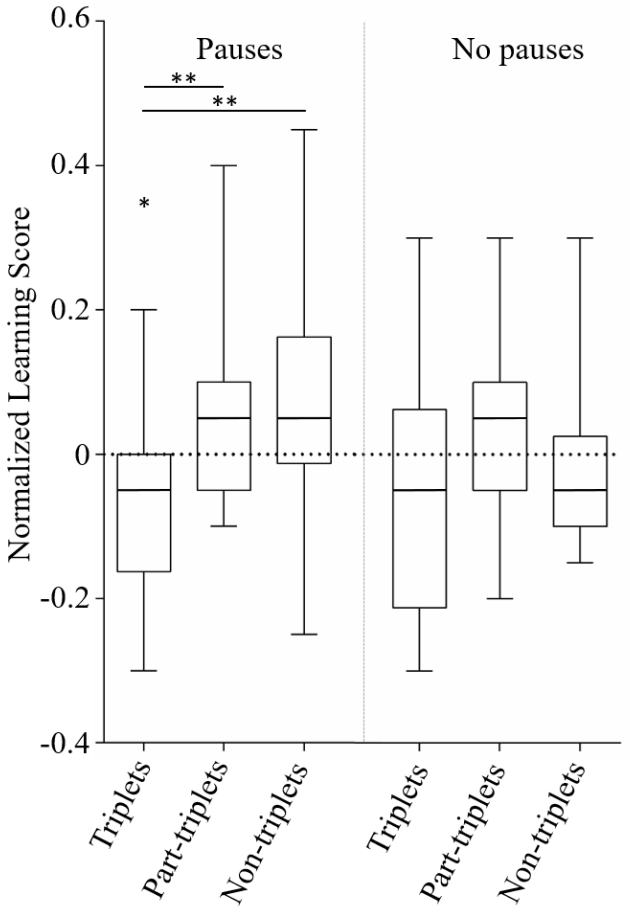


Fig. 3 Learning during the first 20 trials per triplet category. Only when the zebra finches that heard pauses in the first training strings responded differently to triplets of co-occurring elements (consistent triplets) compared to triplets of less co-occurring (partly consistent triplets) or not co-occurring (inconsistent triplets) elements. There was no effect of co-occurrence of elements in the condition without pauses. The boxplots show the median (horizontal line) and first and third quartile of the data, with whiskers extending to the minimum and maximum values. Asterisks with a line indicate a significant difference between the groups, the single asterisk indicates a significant difference from random (=0.5).

These results remained consistent over the first 100 trials of each triplet category. In the Pause-group, the responses to the consistent triplets stayed lower than the responses to the partly consistent and inconsistent triplets (lmm Tukey c-triplets vs. pc-triplets: $z = -2.70$, $P = .02$; c-triplets vs. ic-triplets: $z = -3.14$, $P = .005$; pc-triplets vs. ic-triplets: $z = -0.44$, $P = .90$; Fig. 4). The No-pause group continued to not show a difference in response to any of the triplet categories (lmm Tukey c-triplets vs. pc-triplets: $z = -1.16$, $P = .478$; c-triplets vs. ic-triplets: $z = -0.33$, $P = .94$; pc-triplets vs. ic-triplets: $z = 0.83$, $P = .68$; Fig. 4).

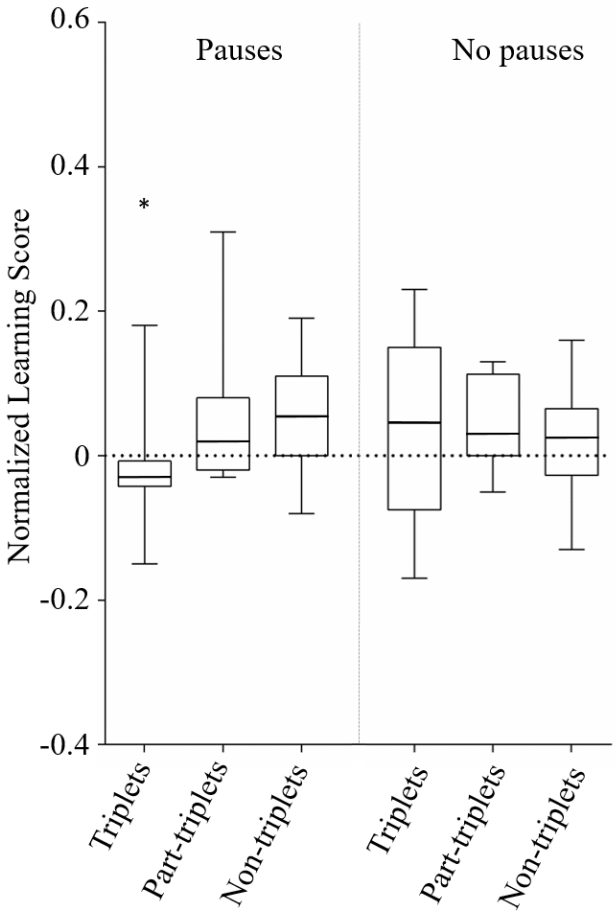


Fig 4. Results of the first 100 trials per triplet category. Zebra finches that heard pauses in the string discrimination training made more incorrect responses to the consistent triplets compared to the partly consistent triplets or inconsistent triplets. There was no difference in response between the triplet categories in the No-pause group. The boxplots show the median (horizontal line) and first and third quartile of the data, with whiskers extending to the minimum and maximum values. Asterisk indicates a significant difference from random (=0.5).

There was no difference in response to one of the two consistent triplets, one of the two partly consistent triplets or one of the two inconsistent triplets, except for the partly consistent triplets in the No-pause group, which responded better to the triplet that had occurred earlier in the string (mean %C 1=0.58, mean %C 2=0.48, P=.01), meaning that in this case the location of the elements in the discrimination string influenced the recognition of this particular segment. No significant correlation was found between the duration of the discrimination training and the results in the segments reversal training (Pearson %C triplets with training days = -0.25).

DISCUSSION

The results of the string discrimination training indicate that zebra finches discriminated more readily between two strings when there were longer pauses between element triplets. It cannot be excluded that this enhanced discrimination is affected by the total duration of the strings, which increases with increased pause length. This means that the birds are exposed to a lower number of elements per unit of time. However, given that longer inter-element pauses in natural songs are more likely to be perceived as a break in a song string, a slower succession of elements seems unlikely to result in a better memory of which elements are to be followed by which others. Rather, and in-line with the finding of segments reversal training, we suggest that the pauses aid in detecting co-occurring element triplets and that this improves learning of strings consisting of such triplets. This interpretation is in-line with the results of many studies showing that strings organized in chunks, from telephone numbers for humans to sequences of visual tokens to be learned by pigeons or rats, are memorized faster and better than strings with an equal number of items providing no chunking cues (e.g., 29-31).

The results of the segments reversal training demonstrate that pauses in strings of song elements elicited an enhanced memorization of co-occurring element triplets that were surrounded by pauses. The mere co-occurrence of song elements in strings without pauses did not evoke a similar response, as demonstrated by the lack of an association between triplet category and %C in the reversal training of the No-pause group. This indicates that pauses between chunks positively affect the memorization of these chunks.

In their natural songs, zebra finches produce longer pauses between and shorter pauses within chunks (32). We suggest that the natural longer inter-chunk pauses enhance song memorization in young birds and might therefore play an important role in the song-learning process (10, 11, 13, 14, 17, 18). Moreover, because our experiment was conducted with adult zebra finches and individuals of both sexes, we show that this proposed learning advantage of the presence of longer inter-chunk pauses is not specific for the period of song production learning. The zebra finches also memorized chunks of co-occurring elements better with longer pauses between such chunks, even though their song learning phase had finished. This suggests that under natural conditions, the presence of such pauses may help adult birds to memorize the songs of different individuals with rather similar songs and hence may help to discriminate among individuals. Bengalese finches, a related vocal learning

species, use either co-occurrence, longer pauses or both when they are copying parts of songs from their tutor (18). In natural songs, as used in the aforementioned experiment, these two factors are strongly correlated. The results from our study show that zebra finches are more sensitive to these pauses than to co-occurrence. This could be an indication that when songbirds have two correlating cues that they can use, like co-occurrence and pauses, they might be more prone to pay attention to the pauses.

Although co-occurrence of elements on its own did not create better recognition of element groups by zebra finches in the present experiment, an earlier study (21) showed that zebra finches can use co-occurrence for sequence discrimination and also remember co-occurring items when co-occurring elements are reshuffled in position within sequences. The birds could use similarity in both transitional and positional information of the training strings to discriminate between new strings. Interestingly, the zebra finches responded to the most reliable cue in that particular context, indicating a context-dependent learning strategy. Apart from songbirds, other animals can also respond to element co-occurrence. Tamarins (Saguinus oedipus) and rats (Rattus rattus), for instance, both respond more strongly to segments of a string with high co-occurrence (33, 34). Neither species are considered to be vocal learners, demonstrating that a tendency to attend to co-occurrences is not specific to language or song learning.

In the string discrimination training of the present study, zebra finches of the No-pause group could have used both transitional (co-occurrence) and positional information to make the discrimination. Knowing, however, that they do not differentiate among the different types of triplets, it is likely that the zebra finches made the discrimination using the position of the elements (similar to some of the birds in 21). This indicates a learning strategy comparable to visual sequence learning in other species (28, 35, 36). In these studies, animals used positional and ordinal information to memorize the sequential organization of items. Moreover, tamarins and rhesus monkeys were able to learn sequences that could not be chunked and responded above chance to all two-item combination from these sequences. This is an indication that, unlike pigeons and rats, monkeys formed an ordinal representation of the sequences (37, 38). Likewise, chunking of longer sequences might be a useful tool in memorizing conspecific songs in zebra finches.

Human infants also tend to use pauses as a cue to find word boundaries (1-4). When pauses are present in a speech stream, infants tend to treat inter-pause segments as more familiar than segments that span pause boundaries. This is quite similar to the

responses of the zebra finches. However, human infants are also able to use co-occurrence or transitional probabilities between elements when pauses are not a reliable cue (5-7). Although zebra finches are able to use transitional probabilities (21) and might use them during song learning (39), the current experiment shows that they do not readily form an association with chunks formed by co-occurrence only.

In conclusion, longer pauses between chunks in strings of song elements aid zebra finches in the song recognition process. These pauses also stimulate memorization of segments of strings that are determined by such pauses. The co-occurrence of song elements on its own does not elicit similar learning advantages. These results indicate that pauses between chunks of song elements might function not only as an aid to song learning in juvenile birds, but also to song discrimination in adult birds of both sexes.

REFERENCES

1. Nazzi T, Nelson DGK, Jusczyk PW, Jusczyk AM (2000) Six-month-olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. Infancy 1(1):123–147. doi:10.1207/S15327078IN0101_11

2. Johnson EK, Jusczyk PW (2001) Word segmentation by 8-month-olds: when speech cues count more than statistics. J Mem Lang 44(4):548–567. doi: 10.1006/jmla.2000.2755

3. Thiessen ED, Saffran JR (2003) When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. Dev Psychol 39(4):706–716. doi:10.1037/0012-1649.39.4.706

4. Lew-Williams C, Pelucchi B, Saffran JR (2011) Isolated words enhance statistical language learning in infancy. Dev Sci 14(6):1323–1329. doi:10.1111/ j.1467-7687.2011.01079.x

5. Saffran JR, Aslin RN, Newport EL (1996a) Statistical learning by 8-month-old infants. Science 274(5294):1926–1928. doi:10. 1126/science. 274.5294.1926

6. Saffran JR, Newport EL, Aslin RN (1996b) Word segmentation: the role of distributional cues. J Mem Lang 35(4):606–621. doi:10. 1006/jmla.1996.0032

7. Aslin RN, Saffran JR, Newport EL (1998) Computation of conditional probability statistics by 8-month-old infants. Psychol Sci 9(4):321–324. doi: 10.1111/1467-9280.00063

8. Brent MR, Cartwright TA (1996) Distributional regularity and phonotactic constraints are useful for segmentation. Cognition 61(1–2):93–125. doi: 10.1016/S0010-0277(96)00719-6

9. Swingley D (2005) Statistical clustering and the contents of the infant vocabulary. Cogn Psychol 50(1):86–132. doi:10.1016/j.cogp sych.2004.06.001

10. Hultsch H, Todt D (1989a) Memorization and reproduction of songs in nightingales (*Luscinia megarhynchos*)—evidence for package formation. J Comp Psychol A 165(2):197–203. doi:10.1007/ BF00619194

11. Hultsch H, Todt D (1989b) Song acquisition and acquisition constraints in the nightingale, *Luscinia megarhynchos*. Naturwissenschaften 76(2):83–85. doi: 10.1007/BF00396717

12. Hultsch H, Todt D (2004) Approaches to the mechanisms of song memorization and singing provide evidence for a procedural memory. An Acad Bras Cienc 76(2):219–230. doi:10.1590/ S0001-37652004000200005

13. ten Cate C, Slater PJB (1991) Song learning in zebra finches—how are elements from 2 tutors integrated. Anim Behav 42:150–152. doi:10.1016/ S0003-3472(05)80617-7

14. Williams H, Staples K (1992) Syllable chunking in zebra finch (Taeniopygia guttate) song. J Comp Psychol 106(3):278–286. doi: 10.1037/0735-7036.106.3.278

15. Cynx J (1990) Experimental determination of a unit of song production in the zebra finch (Taeniopygia guttata). J Comp Psychol 104(1):3–10. doi: 10.1037/0735-7036.104.1.3

16. Franz M, Goller F (2002) Respiratory units of motor production and song imitation in the zebra finch. J Neurobiol 51(2):129–141. doi:10.1002/neu. 10043

17. Suge R, Okanoya K (2010) Perceptual chunking in the self-produced songs of Bengalese finches (Lonchura striata var. domestica). Anim Cogn 13(3):515–523. doi:10.1007/s10071-009-0302-4

18. Takahasi M, Yamada H, Okanoya K (2010) Statistical and prosodic cues for song segmentation learning by Bengalese finches (Lonchura striata var. domestica). Ethology 116(6):481–489. doi:10.1111/j.1439-0310.2010.01772.x

19. Okanoya K (2004) Song syntax in Bengalese finches: proximate and ultimate analyses. Adv Stud Behav 34:297–346. doi:10.1016/S0065-3454(04)34008-8

20. Seki Y, Suzuki K, Takahasi M, Okanoya K (2008) Song motor control organizes acoustic patterns on two levels in Bengalese finches (Lonchura striata var. domestica). J Comp Psychol A 194(6):533–543. doi:10.1007/ s00359-008-0328-0

21. Chen J, ten Cate C (2015) Zebra finches can use positional and transitional cues to distinguish vocal element strings. Behav Process. doi:10.1016/ j.beproc.2014.09.004

22. van Heijningen CA, de Visser J, Zuidema W, ten Cate C (2009) Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. PNAS USA 106:20538–20543. doi:10.1073/pnas. 0908113106

23. van Heijningen CA, Chen J, van Laatum I, van der Hulst B, ten Cate C (2013) Rule learning by zebra finches in an artificial grammar learning task: which rule? Anim Cogn 16:165–175. doi:10.1007/ s10071-012-0559-x

24. Chen J, van Rossum D, ten Cate C (2014) Artificial grammar learning in zebra finches and human adults: XYX vs XXY. Anim Cogn 18:151–164. doi: 10.1007/s10071-014-0786-4

25. ten Cate C (2014) On the phonetic and syntactic processing abilities of birds: from song to speech and artificial grammars. Curr Opin Neurobiol 28:157–164. doi:10.1016/j.conb.2014.07.019

26. Spierings MJ, ten Cate C (2014) Zebra finches are sensitive to the prosodic features of human speech. Proc R Soc B 281(1787):20140480. doi:10.1098/ rspb.2014.0480

27. Mackintosh NJ (1969) Further analysis of the overtraining reversal effect. J Comp Physiol Psychol 67(2):1–18. doi:10.1037/h0026784

28. Boersma P, Weenink D (2014) Praat: doing phonetics by computer [Computer program]. Version 5.3.04, Retrieved 6 Dec 2013 from http://www.praat.org/

29. Terrace HS (1987) Chunking by a pigeon in a serial-learning task. Nature 325(6100):149–151. doi:10.1038/325149a0

30. Terrace HS, Chen S (1991) Chunking during serial learning by a pigeon: II. Integrity of a chunk on a new list. J Exp Psychol Anim B 17(1):94–106. doi: 10.1037/0097-7403.17.1.94

31. Fountain SB, Rowan JD, Muller MD, Kundey SMA, Pickens LRG, Doyle KE (2012) The organization of sequential behaviour. In: Zentall TR, Wasserman EA (ed) The Oxford handbook of comparative cognition. Oxford University Press, Oxford, pp 594–614

32. Zann R (1993) Structure, sequence and evolution of song elements in wild Australian zebra finches. Auk 110(4):702–715. doi:10. 2307/4088626

33. Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. Cognition 78(3):B53–B64. doi:10.1016/ S0010-0277(00)00132-3

34. Toro JM, Trobalon JB (2005) Statistical computations over a speech stream in a rodent. Percept Psychophys 67(5):867–875. doi:10. 3758/BF03193539

35. Brannon EM, Terrace HS (1998) Ordering of the numerosities 1 to 9 by monkeys. Science 282(5389):746–749. doi:10.1126/science. 282.5389.746

36. Fountain SB, Benson DM (2006) Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning. Learn Motiv 37(2):95–112. doi:10.1016/j.lmot.2005.09.002

37. Swartz KB, Chen S, Terrace HS (1991) Serial learning by Rhesus monkeys: I. Acquisition and retention of multiple four-item lists. J Exp Psychol Anim B 17(4):396–410. doi:10.1037/0097-7403.17.4.396

38. Ohshiba N (1997) Memorization of serial items by Japanese monkeys, a chimpanzee, and humans. Jpn Psychol Res 39(3):236–252. dos 10.1111/ 1468-5884.00057

39. Lipkind D, Marcus GF, Bemis DK, Sasahara K, Jacoby N, Takahasi M, Suzuki K, Feher O, Ravbar P, Okanoya K, Tchernichovski O (2013) Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. Nature 498(7452):104–108. doi:10.1038/nature12173

Michelle Spierings, Jeroen Hubert & Carel ten Cate

A different version of this manuscript has been submitted for publication



3

ZEBRA FINCHES GROUP TONES ALTERNATING IN PITCH AS TROCHEES

ABSTRACT

Humans have a strong tendency to spontaneously group visual or auditory stimuli together in larger patterns. One of these perceptual grouping biases is formulated as the Iambic/Trochaic Law, where humans group successive tones alternating in pitch as trochees (initial prominence) and alternating in duration as iambs (final prominence). The grouping of pitch alternations into trochees is found in humans around the globe and in one non-human animal species, rats. The perceptual grouping of sounds alternating in duration is dependent on the participants' native language and has so far not been found among animals. In the current study we explore the extent to which both perceptual biases are shared between humans and a songbird, the zebra finch. The zebra finches were trained to discriminate between short strings of pure tones organized as iambs and pure tones organized as trochees. One group received tones that alternated in duration, the other group heard tones alternating in pitch. After correct discrimination, the zebra finches were exposed to longer ambiguous strings of alternating sounds. The zebra finches categorized ambiguous strings of alternating tones as trochees, similar to humans. However, the zebra finches in the duration condition did not learn to discriminate between training stimuli organized as iambs and trochees. This study shows that the perceptual bias to group tones alternating in pitch as trochees is not specific to humans and rats, but may be a more widespread perceptual bias among animals.

INTRODUCTION

When hearing a long string of successive tones, humans tend to perceive them as a concatenation of duplets that either have prominence on the first tone (trochees) or prominence on the second tone (iambs). When the tones are alternating in pitch or intensity they are grouped as trochees, tones alternating in duration are often, but not universally, grouped as iambs (1-4). This grouping principle according to the Iambic-Trochaic Law (ITL) has been noted already a century ago (5, 6) and has been confirmed by numerous studies ever since (1, 2, 7, 8).

The ITL applies not only to adult listeners, but also infants show perceptual grouping of sound strings. Already from 5 months of age infants show an increased brain response to trochees in a string of iambs and have a preference for the iambic or trochaic stress pattern of their native language (9, 10). In behavioural paradigms this perceptual ability and grouping bias becomes clear around 8 months of age, when English speaking infants segment strings of tones alternating in intensity as trochees and alternating in duration as iambs (11). The early onset of the ITL strengthened the idea that this might be a universal principle that is shared between the different age classes. Another indicator of a universal grouping principle is the fact that the perceptual grouping principles are not restricted to a particular sound type: human adults and infants show perceptual grouping of musical tones, beeps, or spoken syllables (e.g. 3, 8, 12-14). Lastly, there is great similarity between the principles of the ITL and the Gestalt principles applying to perception of visual objects, also vouching for a universal perceptual principle (for a review, see 15).

In language perception, the ITL plays a large role in the perception of words and the segmentation of speech streams (4, 16). For example, in English 90% of the words have a trochaic stress pattern (17). In line with this, English infants are better at recognizing trochees in a string of continuous speech sounds (17, 18) and already have a general preference for listening to trochees over iambs (19). Even learning a second language does not interfere with this perceptual bias (20). Furthermore, when infants are faced with the task to segment a string of speech sounds into words, they use the natural stress pattern of their native language as a clue for word boundaries (21-23). So again, the way the strings are segmented is related to the native language of the listener, French and German infants use different stress patterns to recognize words in longer sound strings (24).

The effect of acoustic experience on perceptual grouping biases is, thus far, only known to be present in the grouping of duration alternations. For example, native speakers of English, German and Spanish group duration alternations as iambs (3, 14, 25, 26), whilst adult Zapotec speakers and Japanese learning infants group these alternations as trochees (8, 27). This effect of experience might indicate that it is only the trochaic grouping of intensity or pitch alternations that is universally shared between languages, ages and domains. This arises the question whether this grouping principle is unique to humans or can also be found among other animals. An indication that it may be a more general perceptual phenomenon is that rats also group tones alternating in pitch as trochees (28, 29). The rats were trained to discriminate tonal strings alternating in pitch or duration from strings in which the tones were randomly organized. They only received food for pressing a lever after hearing the alternating strings. After they learned to discriminate they were exposed to pairs of tones, either iambs or trochees. Their lever presses revealed that the rats grouped the pitch alternating strings as trochees and did not group the tones in the duration alternating strings (28). A follow-up showed that when rats were passively exposed to either iambic or trochaic stress patterns, they would group duration alternating strings in accordance to the pattern they were exposed to (29). Thus, similar to humans, acoustic experience influences the perceptual grouping bias.

The perceptual grouping bias of the rats aids to the suggestion that the trochaic grouping bias is not specific to language or humans and may be an ancient principle that humans use to organize speech sounds (30). However, with no other animal species tested, the generality of the grouping bias is not clear and it might not be shared among a wider range of species.

In the current study we explore the presence of iambic or trochaic grouping biases in a bird species, the zebra finch. Zebra finches, small songbirds, are a well-studied model species for auditory perception (31, 32). Also, they are able to perceive stress in human speech and are sensitive to the stress pattern over a string of speech syllables (33), something that has also been demonstrated in budgerigars (34). This shows that birds are sensitive to acoustic features that also influence the iambic/trochaic grouping bias in humans and make birds an excellent group to examine for the presence and direction of grouping biases.

In our study, we trained zebra finches to discriminate between iambs and trochees constructed of tones in a go left/go right paradigm. These tones varied in pitch for one group of animals and varied in duration for the other group. After the zebra finches correctly discriminated between the iambic and the trochaic structures, they

were tested with long strings of alternating tones, again either alternating in pitch or in duration. If they perceived these alternations as iambic, they were expected to give a similar response as to the trained iambs. If they perceived them as trochaic, they were expected to give a response similar to that of the trochaic training stimuli.

METHODS

Subjects

Sixteen zebra finches were tested (8 males, 8 females) and were distributed equally over two experimental groups. All zebra finches were at least 160 days old at the beginning of the experiment. The animals were bred and reared at the Leiden University animal breeding facility, where they were housed in single sex groups on a 13.5 L: 10.5 D schedule at 20-22 ºC. Food, water, grit and cuttlebone were available *ad libitum*. During the experiment food was used as reinforcement and therefore only available after a correct trial. Food intake was monitored daily and additional food was provided whenever necessary. All experiments were approved by the Leiden committee for animal experimentation DEC number 14229.

Apparatus

The experiments took place in individual operant conditioning cages, which were placed in separate sound attenuated rooms. Each room was illuminated by a fluorescent tube that emitted a daylight spectrum on the same 13.5 L: 10.5 D schedule as was used in the breeding facility. A speaker (Vifa 10BGS119/8) was located 1m above the center of the cage. The operant conditioning cages were constructed of mesh wire sides with a back wall and floor of foamed PVC. The back wall supported three horizontally aligned pecking keys and a food hatch above them, all easily accessible from provided perches. The pecking keys were fitted with red LED lights. Birds needed to peck on the middle key to initiate a trial and stimulus playback. Depending on the nature of the playback, the bird had to either peck on the key on the left or the key on the right within 30 seconds. A correct response was followed by 8 seconds of food access, an incorrect response was followed by 15 seconds of darkness.

Stimuli

Training. The birds were trained to discriminate between stimuli consisting of two duplets with both iambic stress and stimuli of two duplets with trochaic stress. For one group of birds the stress was created by changes in pitch, for the other group by

changes in duration (see figure 1 for an example). The training stimuli were concatenations of four pure tones, two types organized in a ABAB structure. Each bird received a set of four stimuli with iambic stress and four with trochaic stress. In the duration condition, the iambic stimuli had a short-long-short-long structure and the trochaic stimuli had a long-short-long-short concatenation. Each of the four training stimuli within one category (iambs or trochees) started with a different tone duration, the long tones were always 50% longer than the short tones within the same quadruplet. In the pitch condition the stimuli were organized in a similar fashion, the iambic stimuli had a low-high-low-high concatenation and the trochaic stimuli a high-low-high-low concatenation. The high tones were always 25% higher than the low tones within the same quadruplet. The pure tones were always 70 dB and were separated by a 60 msec. silent interval. Per condition (pitch and duration) four different training sets were created to avoid any pseudoreplication (see table 1 for an overview of the training stimuli).

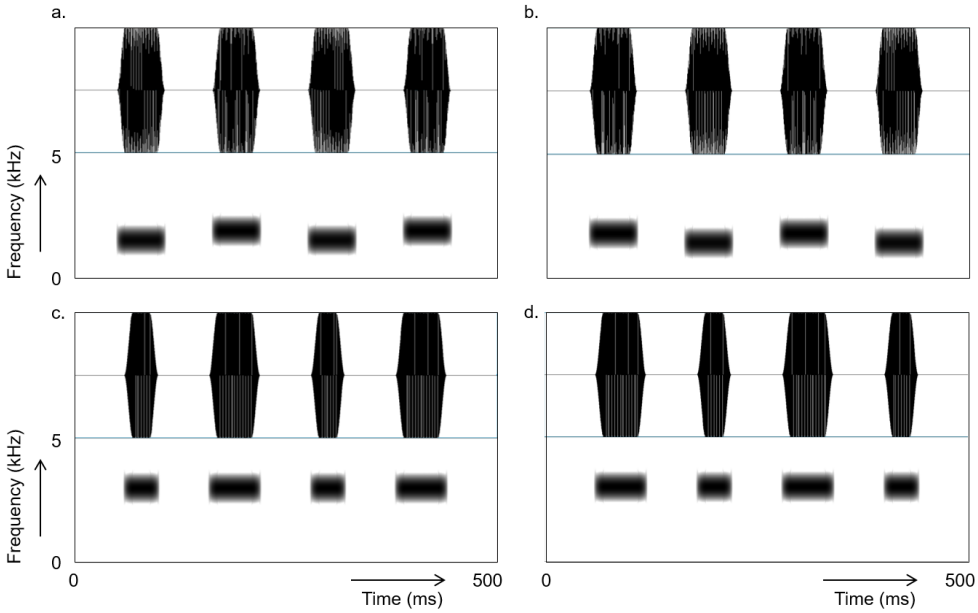


Figure 1. Example of four training stimuli. Two with changes in pitch: one stimulus with two duplets with iambic stress (top left) and one stimulus with two duplets with trochaic stress (top right). And two with changes in duration: duplets with iambic stress (bottom left) and duplets with trochaic stress (bottom right).

Tests. There were four different test conditions, three testing a potential bias of the birds (test 1, 2 and 3) and one control condition (test 4, table 2). In test 1, 2 and 3, the zebra finches heard long sequences of alternating tones. If they considered these to be organized in an iambic way, they were expected to categorize them as they did with the iambic training stimuli. If they grouped the tones as trochees, they should respond similarly as to the trochaic training stimuli. Test 4 also had long strings, but consisting of one single tone. These strings could not be grouped based on the alternations, which means that non-random responses of the birds indicated a response preference for one of the keys or a perceptual grouping bias extended to non-alternating sounds.

More specifically, test 1 consisted of the same tones that were used for the training stimuli, only now in a 26 tones long concatenation. As in the training stimuli, the tones were alternating with long-short or high-low configurations. However, unlike the training stimuli, these long strings started and ended with a 1.3 sec. fade, making it difficult to determine with which tone the string started. Moreover, the birds received both test strings starting with a stressed tone as well as test strings starting with an unstressed tone. All tones were 70 dB and were separated by 60 ms silent intervals. Test 2 and 3 were constructed similarly to test 1, but now consisted of new tones. These tones were either of a different duration than the tones from the training stimuli (in the duration condition) or had a different pitch than the training tones (in the pitch condition). Test 2 had tones within the range of the training tones, test 3 consisted of tones that were higher and lower, or longer and shorter than the training tones. Just like test 1, these strings were 26 tones long and had a 1.3 sec fade in and fade out. Test 4 consisted of three different test strings, all containing one tone that also occurred in the training strings. These test strings were therefore no alternation of high and low or long and short tones, but a repetition of a single tone. Just like the strings of test 1, 2 and 3, they were 26 tones long, separated by 60 msec silent intervals and with a fade in and fade out of 1.3 seconds. As there were four different training sets created for each condition, there were also four test sets per condition to match the stimuli from the training.

Table 1. Overview of the duration or pitch training stimuli for the duration condition (a) and the pitch condition (b). The table shows the values of the tones used in the training stimuli rounded to the nearest integer. Shown here is one of the sets used, starting with 40 msec (duration condition) and 1500 Hz (pitch condition). The other three sets were created with a longer or higher start tone start tone (45, 50 and 55 msec. and 1750, 2000 and 2250 Hz). The relative difference between two consecutive tones remained constant.

a.

Duration training	**Iambic stimuli** (Tone durations in msec)	**Trochaic stimuli** (Tone durations in msec)
1	40-60-40-60	60-40-60-40
2	60-90-60-90	90-60-90-60
3	90-135-90-135	135-90-135-90
4	135-202-135-202	202-135-202-135

b.

Pitch training	**Iambic stimuli** (Tone frequencies in Hz)	**Trochaic stimuli** (Tone frequencies in Hz)
1	1500-1875-1500-1875	1875-1500-1875-1500
2	1875-2344-1875-2344	2344-1875-2344-1875
3	2344-2930-2344-2930	2930-2344-2930-2344
4	2930-3662-2930-3662	3662-2930-3662-2930

Table 2. Overview of the test stimuli. All test strings were 26 tones long and had a fade in and fade out of 1.3 sec. Test strings were presented in 20% of the trials when the zebra finch had reached the standard training criterion.

	Tones in stimuli	**Alternating high/low or short/long**	**Number of different strings**
Test 1	Same as training tones	Yes	4
Test 2	New, within range of training tones	Yes	2
Test 3	New, outside range of training tones	Yes	4
Test 4	Same as training tones	No	6

Experimental design

Each zebra finch was first trained on the go left/go right design with two unfamiliar zebra finch songs. They received a reward for pecking on the left key after hearing one song, and on the right key after hearing the other song. When they reached the standard criterion of over 75% correct responses to both songs for three consecutive days, they proceeded to the training.

During training the zebra finches had to discriminate between four stimuli with iambic stress and four stimuli with trochaic stress by pecking on either the left or the right key after the stimulus was played. For half of the birds the key for iambs was on the right side of the cage and the key for trochees on the left, for the other half of the birds this was switched. If an individual only used one of the response keys instead of both, the program was set to repeat a stimulus that received an incorrect response until the bird gave the correct response. This setting would be on for less than 24 hours, motivating the animal to use both response keys. Training continued until the birds reached the standard criterion or when they reached 20,000 trials without having 3 consecutive days with more than 55% correct responses. Only when they reached the learning criterion would they proceed to the test phase.

In the test phase 20% of the trials were non-reinforced test stimuli, presented in a random order within a test block. The test items were organized in two sequentially presented test blocks, one with the stimuli of test 1, 2 and 3, the second one with the stimuli of test 4. A bird continued to the next test block when each test stimulus in the block was presented 40 times.

Analysis

The responses of the birds to the training and test stimuli were calculated as proportions of responses to the iambic and the trochaic key per stimulus (number of responses/number of trials). The birds could also not respond after initiating a trial, which was recorded as the fraction of no responses (times not responded/number of trials). These three fractions, iambic, trochaic and no response, always add up to be a hundred percent per stimulus. An average response fraction per test was calculated by taking the average fractions towards the different stimuli within one test condition. Also for the training, we calculated the average responses to all training iambic training stimuli and the average response to the trochaic stimuli per bird. These data were analyzed with a generalized linear model (glm) with test item (all tests and the training iambic and trochaic stimuli) as fixed effect and the individual as the random measure.

Pairwise comparisons were made between the fractions of responses to the iambic and the trochaic key for each test and the two training sets by using a Tukey's post-hoc test, corrected for multiple testing.

RESULTS

Training

The training of the birds lasted until they reached the standard discrimination score of over 75% correct responses to both iambs and trochees, or until they did 20,000 trials. All birds in the pitch condition learned the discrimination in less than 20,000 trials with an average of 14,717 trials (+/- 4118). None of the birds in the duration condition were able to learn the discrimination within the 20,000 trial frame (both conditions shown in figure 2).

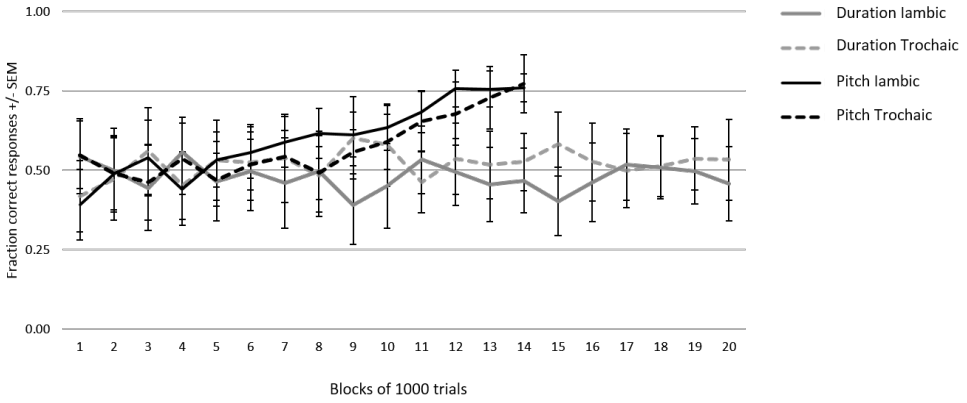


Figure 2. Proportions of correct responses to the iambic and the trochaic training sounds. Dur Iambic are quadruplets with increased duration of the second and fourth tone. Dur trochaic are quadruplets with increased duration of the first and third tone. In the same fashion, Pitch iambic are quadruplets with increased frequency of the second and fourth tone and Pitch trochaic are quadruplets with increased frequency on the first and third tone. The lines show the average responses of the 8 zebra finches in each condition (duration and pitch) organized in blocks of 1000 trials.

Test

Only the zebra finches in the pitch condition could be tested, as these were the only individuals reaching the standard discrimination criterion during training. In the pitch test, the zebra finches responded to the stimuli of test 1, long alternating strings of known tones, by pecking more often on the trochaic key than on the iambic key (mean

iambic=0.19, mean trochaic=0.43, $p<0.01$, figure 3). In 38% of the trials no pecking response was given. A similar result was found for the strings of test 2, long alternating strings with new tones within the range of the training tones. The zebra finches responded by pecking on the trochaic key more often than on the iambic key (mean iambic=0.23, mean trochaic=0.42, $p<0.01$). The zebra finches did not respond in 35% of the trials. When the tones in the long strings were outside the training range (test 3) the zebra finches did not peck more on the iambic or the trochaic key (mean iambic=0.2, mean trochaic=0.21, $p=0.87$). The zebra finches responded less to the stimuli of this test, with no pecking response in 58% of the trials. In test 4, strings without the high-low alternation, the zebra finches also responded equally often by pecking on the iambic as on the trochaic key (mean iambic=0.23, mean trochaic=0.25, $p=0.64$). In all test conditions the birds did not show a different response to the test strings that started with a low tone and test strings that started with a high tone (all $p>0.1$).

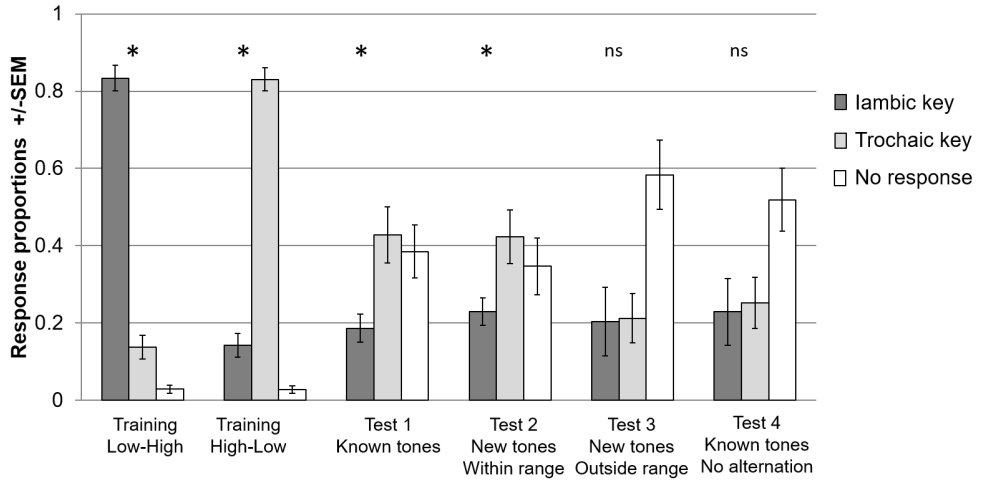


Figure 3. Proportions of responses to the training and test stimuli of the pitch condition. The dark grey bars show the proportions of pecks on the iambic key, the light grey bars shown the pecks on the trochaic key. The white bars show the proportion of trials to which the birds did not respond by pecking on a key. The bars shown the averages of all 8 zebra finches, the error bars show the SEM.

DISCUSSION

All zebra finches in this study learned to discriminate between trochees and iambs when the tones varied in pitch. Sequential tests showed that the zebra finches had grouped the tones of long alternating strings into smaller sets. The birds responded

more to the long strings as being similar to the trochees than to the iambs of the training, which shows a trochaic grouping of pitch alternating tones, similar to humans. The zebra finches did not have a grouping bias for strings with a repetition of one tone, without pitch alternations. In comparable tasks, humans do group non-alternating tone strings as containing iambs or trochees (3). Moreover, when the tones varied in duration, the zebra finches did not learn to distinguish trochees from iambs within 20,000 trials.

In earlier experiments humans grouped sound string alternating in pitch or intensity into trochees, regardless of the precise nature or familiarity of the sounds (3, 10, 11). Hay and Diehl (3) tested whether adult listeners responded differently to strings consisting of non-speech tones or of synthetic speech sounds. In their experiments the participants (English speaking) grouped strings alternating in intensity or frequency as trochees, regardless of the type of sounds. This general grouping bias is not restricted to the acoustic domain, also in the visual domain do people group alternating objects as trochees (35). For example, when participants saw a string of visual objects that were alternating in flashing rate or in brightness, they grouped them as trochees, the "stressed" objects with a higher flashing rate or brighter objects formed the start of the memorized duplets. This shows that the tendency to perceive long sequences into smaller units with initial prominence is a general mechanism in humans. Grouping of duration alternations, however, is not universal but depends on previous acoustic experience, both in humans (8, 13) as well as in rats (29). The mechanism underlying the trochaic bias seems to be an experience-independent, universally shared mechanism. Our results strengthen this claim, as we show that the trochaic grouping bias is also present in a non-mammal species, the zebra finch. Like humans, these phylogenetically distant animals group tonal strings with frequency alternations as trochees, suggesting that there might be more ancient evolutionary roots to this perceptual mechanism.

The zebra finches did not show perceptual grouping of strings with tones that were outside of their training range. This could be either an effect of the novelty of these tones, or of them being at the limits of the birds' hearing range. In previous acoustic perception studies with zebra finches in a go/no-go paradigm, the animals showed reduced response rates with novel stimuli (36, 37). It is likely that this is due to an avoidance strategy. There are always known training stimuli presented intermixed with the novel test items, making it possible to avoid punishment by not responding to the novel items, while still getting food for correct responses to training items. In our results we see that the zebra finches in general respond less often when they hear a test stimulus. Moreover, when the tones are outside the trained range or when the

tones are not alternating anymore their response rates drop even further. This shows that these strings are probably considered more novel than the long strings with known tones or tones within the training range.

Different bird species, like zebra finches and budgerigars, are known to be sensitive to the prosodic features of human speech (33, 34). Zebra finches can learn to discriminate between quadruplets of speech syllables with initial or final stress created by increasing the pitch, duration and amplitude of a single syllable. This discrimination holds even when only the pitch or the duration cue is increased in the sound. Surprisingly, in the current study the zebra finches were unable to learn to discriminate between tonal quadruplets differing in the ordinal position of long and short tones. It might be that differences in tone durations are only well perceived by zebra finches when they are accompanied by other prosodic cues. Otherwise the differences might have been too subtle, although zebra finches have been shown to perceive these differences in duration (38, 39).

To summarize, the perceptual bias to group pitch variations into iambs is not specific to language or to humans. After it being shown for rats, we now show that zebra finches share the same perceptual primitive. This is a clear indication that the trochaic grouping bias might be a universal perceptual principle and might have been present in our pre-linguistic ancestors.

Acknowledgements

We would like to thank Marisa Hoeschele and Juan Manuel Toro for helpful discussions concerning the design of this experiment. This study was supported by NWO-GW, grant 360.70.452.

REFERENCES

1. Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. In Proceedings of the Thirteenth meeting of the Berkeley Linguistics Society, ed. by M. Niepokuj, et. al., 429–446. Berkeley, CA: Berkeley Linguistics Society.

2. Hayes, B. (1995). Metrical stress theory: Principles and case studies. Chicago: Univ. of Chicago Press.

3. Hay, J.F. & Diehl, R. L. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. Perception & Psychophysics, 69(1), 113-122.

4. Bion, R. H., Benavides-Varela, S. & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. Language and Speech, 54(Pt 1), 123–140.

5. Bolton, T. (1894). Rhythm. American Journal of Psychology, 6, 145–238.

6. Woodrow, H. (1909). A quantitative study of rhythm: The effect of variations in intensity, rate, and duration. Archives of Psychology, 14, 1-66.

7. Vos, P. 1977. "Temporal duration factors in the perception of auditory rhythmic patterns," Scientific Aesthetics/Sciences de l'Art, 1, 183–199.

8. Crowhurst, M. J. & Olivares, A. T. (2014). Beyond the Iambic-Trochaic Law: the joint influence of duration and intensity on the perception of rhythmic speech. Phonology,31(1), 51-94.

9. Weber, C., Hahne, A., Friedrich, M. & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. Cognitive Brain Research, 18, 149–161.

10. Friederici, A., Friedrich, M., and Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. Current Biology, 17, 1208–1211.

11. Trainor, L. J. & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. Perception & Psychophysics, 62, 333–340.

12. Monahan, C.B. & Carterette, E.C. (1985). Pitch and duration as determinants of musical space. Music Perception, 3, 1-32.

13. Iversen, J. R., Patel, A. D. & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. The Journal of the Acoustical Society of America, 124(October), 2263–2271.

14. Bhatara, A., Boll-Avetisyan, N., Agua, T., Hohlesand, B. & Nazzi, T. (2015). Language experience affects grouping of musical instrument sounds. Cognitive science, 1-15.

15. Wagemans. J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M. & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. Psychol. Bull. 138(6), 1172-1217.

16. Hay, J.F. & Saffran, J.R. (2012). Rhythmic grouping biases constrain infant statistical learning. Infancy, 17(6), 610-641.

17. Cutler, A. & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. Comput. Speech Lang. 2, 133 – 142.

18. Jusczyk, P. W., Houston, D. M. & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. Cognitive Psychology, 39, 159–207.

19. Jusczyk, P., Cutler, A. & Redanz, L. (1993). Preference for the predominant stress patterns of English words. Child Development, 64, 675–687.

20. Bijeljac-Babic, R., Hohle, B. & Nazzi, T. (2016). Early prosodic acquisition in bilingual infants: The case of the perceptual trochaic bias. Front. Psychol. 7:210.

21. Morgan, J.L. (1996). A Rhythmic Bias in Preverbal Speech Segmentation. Journal of memory and language, 35, 666-688.

22. Thiessen, E. & Saffran, J. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. Developmental Psychology, 39, 706–716.

23. Thiessen, E. & Saffran, J. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. Language learning and development, 3(1), 73-100.

24. Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J. & Nazzi, T. (2009). The development of language specific prosodic preferences during the first half year of life: evidence from German and French. Infant Behav. Dev. 2, 262–274.

25. Segal, O. & Kishon-Rabin, L. (2012). Evidence for Language-Specific Influence on the Preference of Stress Patterns in Infants Learning an Iambic Language (Hebrew). Journal of speech, language and hearing research, 55, 1329-1341.

26. Crowhurst, M. J. (2013). Perceptual influences on rhythmic grouping preferences. Paper presented at M@90: a workshop on metrical structure in honor of Morris Halle's 90th birthday, MIT.

27. Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., & Werker, J. F. (2010). The development of perceptual grouping biases in infancy: A Japanese-English cross-linguistic study. Cognition, 115(2), 356–361.

28. de la Mora, D. M., Nespor, M. & Toro, J. M. (2013). Do humans and nonhuman animals share the grouping principles of the iambic-trochaic law? Attention, Perception & Psychophysics, 75, 92–100.

29. Toro, J.M. & Nespor, M. (2015). Experience-dependent emergence of a grouping bias. Biology Letters, 11: 20150374.

30. Toro, J.M. (2016). Something old, something new: Combining mechanisms during language acquisition. Current Directions in Psychological Science, 25(2), 130-134.

31. Kriengwatana, B., Escudero, P. & ten Cate, C. (2014). Revisiting vocal perception in non-human animals: a review of vowel discrimination, speaker voice recognition, and speaker normalization. Frontiers in Psychology, 5:1543.

32. Dent, M.L., Martin, A.K., Flaherty, M.M. & Neilans, E.G. (2016). Cues for auditory stream segregation of birdsong in budgerigars and zebra finches: Effects of location, timing, amplitude, and frequency. Journal of the Acoustical Society of America, 139(2), 674-683.

33. Spierings, M.J. & ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. Proc. R. Soc. B., 281, 20140480

34. Hoeschele, M. & Fitch, T. (2016). Phonological perception by birds: budgerigars can perceive lexical stress. Animal Cognition, 19(3), 643-654.

35. Peña, M., Bion, R. A. H., & Nespor, M. (2011). How Modality Specific Is the Iambic–Trochaic Law? Evidence From Vision. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(5), 1199-1208.

36. van Heijningen, C.A., Chen, J., van Laatum, I., van der Hulst, B. & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? Animal Cognition 16:165–175.

37. Chen, J., van Rossum, D. & ten Cate, C. (2015) Artificial grammar learning in zebra finches and human adults: XYX versus XXY. Animal Cognition, 18(1), 151-164.

38. Okanoya, K., & Dooling, R.J. (1990). Detection of gaps in noise by budgerigars (Melopsittacus undulatus) and zebra finches (Peophila guttata). Hear. Res. 50, 185– 192.

39. Nagel, K.I., McLendon, H.M. & Doupe, A.J. (2010). Differential influence of frequency, timing, and intensity cues in a complex acoustic categorization task. Journal of Neurophysiology, 104(3), 1426-1437.

Michelle Spierings & Carel ten Cate

Published in Proceedings of the Royal Society B: Biological Sciences, 2014, 281(1787), 20140480



4

ZEBRA FINCHES ARE SENSITIVE TO THE PROSODIC FEATURES OF HUMAN SPEECH

ABSTRACT

Variation in pitch, amplitude and rhythm adds crucial paralinguistic information to human speech. Such prosodic cues can reveal information about the meaning or emphasis of a sentence or the emotional state of the speaker. To examine the hypothesis that sensitivity to prosodic cues is language independent and not human specific, we tested prosody perception in a controlled experiment with zebra finches. Using a go/no-go procedure, subjects were trained to discriminate between speech syllables arranged in XYXY patterns with prosodic stress on the first syllable and XXYY patterns with prosodic stress on the final syllable. To systematically determine the salience of the various prosodic cues (pitch, duration and amplitude) to the zebra finches, they were subjected to five tests with different combinations of these cues. The zebra finches generalized the prosodic pattern to sequences that consisted of new syllables and used prosodic features over structural ones to discriminate between stimuli. This strong sensitivity to the prosodic pattern was maintained when only a single prosodic cue was available. The change in pitch was treated as more salient than changes in the other prosodic features. These results show that zebra finches are sensitive to the same prosodic cues known to affect human speech perception.

INTRODUCTION

Linguistic communication is made possible by our ability to produce and understand speech utterances. Essential to understanding the meaning of these utterances is sensitivity to paralinguistic information, which is provided by varying parameters like pitch, amplitude and duration of speech segments. These prosodic cues can alter the meaning and emphasis of a sentence and can reveal information about the emotional state of the speaker [1]. They are also important in the process of language acquisition and discrimination [2–6]. Interestingly, we can find variation in similar features when we look at animal vocalizations, like bird song. Zebra finches, for instance, alter the amplitude of their song depending on the distance to the receiver [7] and use different singing speeds for directed and undirected song [8]. Other songbird species can change prosody-like features in their songs to adapt to changing environments [9], or when displaying aggressive behaviour [10].

This shared use and importance of prosodic features in the vocalizations of humans and non-human animals might indicate that they also share the presence of general perceptual sensitivity to prosodic features. There is, indeed, evidence that some songbirds are sensitive to variation in pitch, amplitude and song syllable duration in their natural song [11]. Also, some mammal and bird species show a sensitivity to the general prosodic patterns in human speech [12–16]. These findings suggest a similar sensitivity for prosodic cues in humans and non-human animals. However, the experiments done thus far did not track which features of human prosody were used to make this discrimination. In this study, we used a songbird species, the zebra finch, to examine the sensitivity to various prosodic features in human speech in a more systematic way.

Songbirds are one of the most relevant groups in comparative language and speech research. Like speech, birdsong is characterized by a rapid production of acoustically varying syllables. Unlike the vocalizations in many other groups of animals, bird songs are learned from a tutor and, when acquiring their song, many songbird species go through similar phases to human infants learning language [17,18]. Songbirds are known to exhibit sensitivities to frequency, amplitude and pauses in their own song repertoire [11,19,20]. An indication that they are also sensitive to stress patterns (that may involve similar parameters) in human speech comes from a study on Java sparrows [16]. However, which cues they rely on when listening to human speech, remains an open question.

For our experiments, we tested zebra finches, which are known to be able to discriminate and categorize monosyllabic words and are sensitive to differences in vowel formants in human speech [21–23]. This discrimination holds when the words are uttered by novel speakers or by speakers of a different sex [21]. We arranged human speech syllables to follow an XYXY or an XXYY structure, knowing that zebra finches are able to discriminate between these two structures [24]. Prosodic patterns were added and varied in salience (three, two, one, or no cues) depending on the experiment.

With these experiments we address not only whether zebra finches are capable of detecting and processing prosodic cues in human speech, but also whether the three parameters involved -pitch, amplitude and duration- are all contributing to the perceiving of the prosodic patterns.

MATERIAL AND METHODS

Subjects

Eight zebra finches (four females and four males) from the Leiden University breeding stock were tested in all experiments sequentially. All birds were between 120 and 250 days post hatching and were naive to any other experiments. The birds lived in single sex groups on a 13.5 L:10.5 D schedule at 20–22°C. During the experiment water, grit and cuttlebone were available ad libitum. Food was used as reinforcement and therefore only available after a correct trial. Food intake was monitored daily and the birds received additional food when necessary.

Apparatus

All zebra finch experiments were conducted in an operant conditioning cage (70 (l) 30 (d) 45 (h) cm). Each operant cage was in a separate sound attenuated chamber and illuminated by a fluorescent tube that emitted a daylight spectrum on a 13.5 L : 10.5 D schedule. A speaker (Vifa MG10SD109-08) was located 1 m above the cage. The cage was made from wire mesh except for the floor and a plywood back wall which supported two pecking keys with LED lights. A food hatch was located in between these two keys, easily accessible to the birds. Pecking the left key (sensor 1) elicited a stimulus and illuminated the LED light of the key on the right (sensor 2). Depending on the sound, the bird had to peck sensor 2 or had to withhold its response. A correct response resulted in access to food for 10 s and an incorrect response led to 15 s of darkness.

Stimuli

The stimuli were constructed from eight naturally spoken syllables which were all produced by a male as well as a female speaker. They were first equalized in pitch, amplitude and duration with Praat [25]. The syllables were chosen to contain the same number of heterorganic and homorganic sounds (mo, ka, pu, le, do, sa, nu, fi). Quadruplets were created that followed either an XYXY or an XXYY pattern and were consistent in voice type (male or female). Prosodic patterns, based on natural prosody in human speech [26], were added by modulating the pitch contour, amplitude and duration of the first or last syllable of a quadruplet, creating stressed (with all prosodic cues changed) and unstressed (no extra prosody added) syllables (table 1 and figure 1). The prosodic features were chosen in such a way that the prosodic pattern as well as the syntactic structure were of comparable salience to human adults, as assessed in pilot experiments. When trained with these stimuli, adults readily discriminated between the Xyxy and xxyY structures (capital letters indicate a stressed syllable, lower case an unstressed one). When subsequently asked to discriminate between structures in which the prosody was now switched to oppose the syntactic structure (xyxY and Xxyy), 12 out of the 32 participants based their responses on the prosodic pattern and 16 followed the syntactic structure (four participants did not choose a specific strategy). This was a clear indication that both the prosody and the structure were notable and salient cues for human adults.

Table 1. The parameters for the prosodic cues that were altered to create stressed and unstressed syllables. (The pitch contours describe the change in pitch over the stressed or unstressed syllable: the first number is the frequency at the start of the syllable, the last number is the frequency at the end of the syllable.)

	Unstressed Syllable	Stressed Syllable
Incline/decline pitch contour (Hz)	Female voice 220-190 Male voice 150-120	Female voice 215-250 Male voice 145-180
Amplitude (dB)	70	78
Duration (ms)	250	325

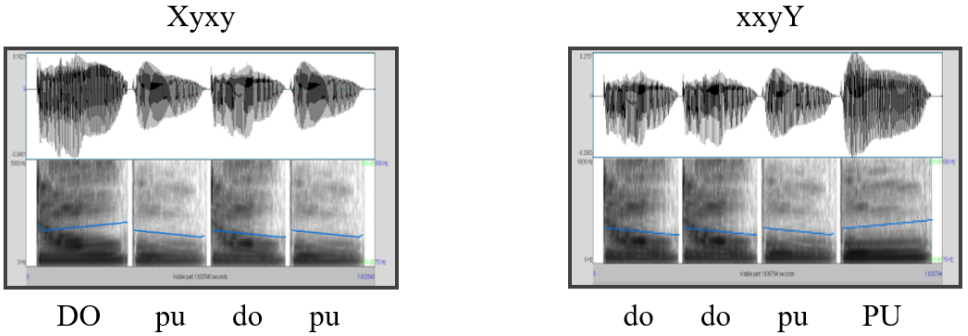


Figure 1. (a,b) Example of two training stimuli. 'DO pu do pu' follows the Xyxy structure and acts as a go-stimulus. The 'do do pu PU' stimulus follows the xxyY structure and is one of the no-go stimuli. In the sonogram, the change in pitch, amplitude and duration of the first or last syllable is visible.

Procedure

Zebra finches were first trained with a zebra finch song as the go stimulus and a pure tone as the no-go stimulus. In this shaping phase, they were familiarized with the go/no-go procedure without exposure to the experimental stimuli. The birds started with, on average, 150 trials a day. Within 3 – 6 days, this number increased to an average of 400 trials a day. During these days, their discrimination performance also increased significantly. When their performance reached our standard criterion (response to positive sounds more than 0.75, response to negative sounds less than 0.25 for 2 successive days), they progressed to the training phase.

During training all subjects were trained with Xyxy quadruplets (prosodic stress on first syllable) as the go stimuli and xxyY quadruplets (prosodic stress on final syllable) as the no-go stimuli. They were presented with three different go-stimuli; Abab, Cdcd and Efef, and three no-go stimuli; aabB, ccdD and eefF. The letters symbolize different syllables, which were also different for each bird. For example, sound A could be 'le' for one bird, but 'fi' for another. Capitalized letters represent a stressed syllable, and small letters an unstressed syllable. When a subject reached the standard criterion, it progressed to the test phase.

The test phase consisted of 80% reinforced training stimuli and 20% non-reinforced test stimuli. The test phase lasted until all the birds had completed 40 trials per stimulus. Experiment 1 tested whether the birds responded more to the prosodic pattern or the syntactic structure of the stimuli by changing these cues in the test phase. Experiment 2 tested whether the birds could generalize the prosodic pattern to test items constructed of new syllables. In experiment 3, we tested whether the birds

would follow the prosodic pattern when only a single prosodic cue was available. Experiment 4 tested the response of the birds to contradicting prosodic cues; one cue was on one edge of the quadruplet and the other two prosodic cues on the opposite edge. The fifth and final experiment tested whether the birds could also discriminate between the syntactic structures when there was no prosodic pattern added. The experiments were done in a stepwise manner, systematically varying the prosodic pattern of the test stimuli. Appendix A provides a full list of stimuli per experiment.

Analysis

For each experiment, we calculated the discrimination ratio (DR) for each stimulus pair and bird as the number of responses to the stimuli that was consistent with the positively reinforced prosodic pattern of the training, divided by the response total. The group data was then analysed with a linear mixed model (lmer) that took into account the repeated measurements per bird and the distribution of the pecking data. Random variables such as individual, sex and type of voice were checked for any influence on the differences in performance. The within-experiment analysis were conducted with a Wilcoxon signed-rank test with continuity correction on the response rates to the stimuli that followed the positively reinforced prosodic pattern from the training and the stimuli that followed the positively reinforced syntactic structure from the training (S^+ = number of correct responses to prosodic pattern stimulus/total number of responses; S^- = number of incorrect responses to syntactic structure stimulus/total number of responses). Performance on an individual level was checked with a Clopper–Pearson binominal confidence interval measurement, using the responses to the positive stimulus and the responses to the negative stimulus per stimulus pair. These statistics are only presented when the birds differed from the group performance.

Experiment 1: prosody versus structure

Stimuli. The test stimuli for this experiment were similar to the training stimuli, but the prosodic pattern and syntactic structure were now switched. The test items were xyxY and Xxyy, while training items remained Xyxy and xxyY. All prosodic cues were kept as explained in the section *Stimuli.*

Results and discussion. Results of experiment 1 showed that zebra finches responded more strongly to the prosodic pattern than to the syntactic structure (mean response Xxyy = 0.89, mean response xyxY = 0.14, Z = 22.53, p = 0.011; figure 2). These results were not influenced by sex of the bird or stimulus voice (sex~test * DR: t =

24.32, p = 0.14; voice~test * DR: t = 25.71, p = 0.27) and were consistent for all individuals (from now on only reported when deviating). During training, the birds could have discriminated the quadruplets on the prosodic pattern, the syntactic structure or both. Our results indicate that in this experiment, all birds focused on the prosodic pattern.

Experiment 2: generalization to unfamiliar syllables

Stimuli. Generalization was measured by changing the syllables from the training quadruplets to new syllables. During training, every bird heard six out of the eight possible syllables. The remaining two syllables were used in this experiment and were therefore different for each bird. The prosodic pattern and syntactic structure of the quadruplets remained the same as the training quadruplets, Xyxy and xxyY.

Results and discussion. Results showed that the zebra finches still discriminated the stimuli (mean S^+ = 0.29, mean S^- = 0.1, Z = 22.524, p = 0.012; figure 2). This could be obtained by either attending to the prosodic pattern or to the syntactic structure of the test items. From the strong response of the zebra finches to the prosodic pattern in experiment 1 and supported by the findings of the subsequent experiments, we conclude that it is most likely that they generalized the prosodic pattern rather than the structural one.

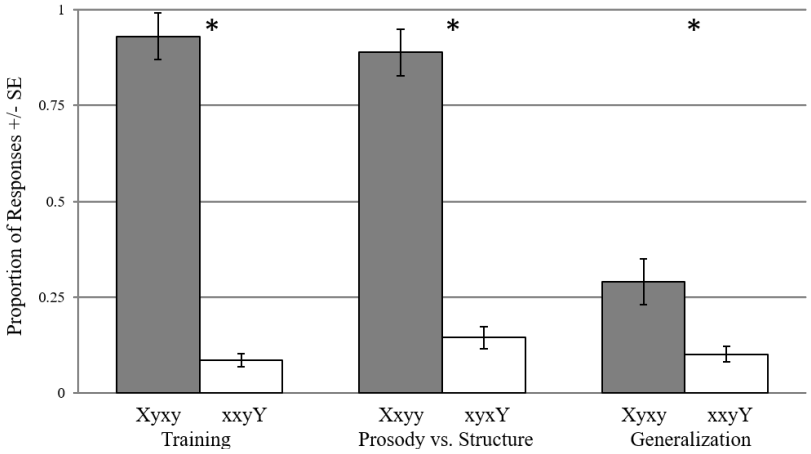


Figure 2. Proportion of responses to the training and test items of experiment 1 and 2. The bars 'training' show the responses of the birds on the day before they started the test. In experiment 1, prosody versus structure, the grey bar indicates a stronger response to the prosodic pattern compared with the syntactic structure. The results of experiment 2, generalization, show responses to similar quadruplets as the training, but consisting of new speech syllables. We show the mean proportion of response of all birds +s.e. The asterisks indicate a significant difference between the responses to the two stimuli types.

Experiment 3: one prosodic cue

Stimuli. To understand the influence of single prosodic cues, the stimuli in this experiment had a single prosodic cue on the first or the last syllable, which could be pitch, amplitude or duration. The syntactic structure remained the same as during training, but the one prosodic cue was switched to the opposite position compared with the training stimuli. Training stimuli remained Xyxy and xxyY, the test stimuli were, for example, xyxY^{pitch} and X^{pitch}xyy.

Results and discussion. Even with one prosodic cue, the birds responded more strongly to the prosodic pattern than to the syntactic structure of the test stimuli (duration: mean responses Xxyy = 0.29, mean responses xyxY = 0.11, Z = 22.524, p = 0.012; amplitude: mean responses Xxyy = 0.43, mean responses xyxY = 0.13, Z = 22.524, p = 0.012; pitch: mean responses Xxyy = 0.42, mean responses xyxY = 0.12, Z = 22.521, p = 0.012; figure 3). There was no difference in DR between the three test conditions (mean DR amplitude = 0.73, mean DR duration = 0.78, mean DR pitch=0.79, t=23.46, p=0.38). This indicates that even a single and less notable prosodic cue is salient enough to outweigh the syntactic structure.

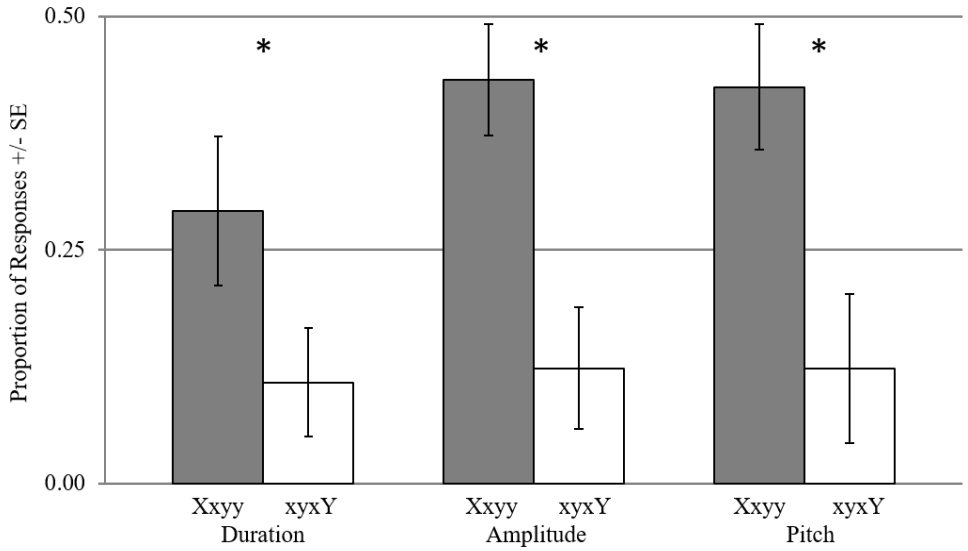


Figure 3. Proportion of responses to test items which had one prosodic cue that was on the opposite position compared with the training stimuli. Shown are the mean proportions of response of all birds +s.e. The asterisks indicate a significant difference between the responses to the two stimuli types. Higher grey bars indicate a stronger response to the single prosodic cue compared with the syntactic structure.

Experiment 4: two prosodic cues versus one prosodic cue

Stimuli. Stimuli for this experiment were constructed with all prosodic cues, but not on the same syllable. Either the first or the last syllable of a quadruplet had two prosodic cues stressed (e.g. pitch and duration); the other syllable was stressed by the one remaining prosodic cue (e.g. amplitude). The quadruplets were ordered such that the two prosodic cues were contradicting the syntactic structure from the training. The single cue was in line with the position that prosody had on the syntactic training structure. When trained on Xyxy and xxyY, a test quadruplet looked like $X^{pitch,duration}$ $xyy^{amplitude}$ and $x^{amplitude}yxY^{pitch,duration}$. Here, the syllable that has one prosodic cue is indicated by a small, underlined letter.

Results and discussion. Results showed that when an increased pitch on a syllable was combined with any of the other two prosodic cues, this was salient enough for the birds to respond to it as if it were the training prosody (A and P versus D: mean responses Xxyy = 0.51, mean responses xyxY = 0.14, Z = 22.521, p = 0.012; D and P versus A: mean responses Xxyy = 0.58, mean responses xyxY = 0.17, Z = 22.380, p = 0.017; figure 4), even though a third cue was different from this pattern and corresponded to the syntactic structure. However, when an increased pitch was the single cue, the birds did not discriminate between the two structures (A and D versus P: mean responses Xxyy = 0.36, mean responses xyxY = 0.24, Z = 20.840, p = 0.401; figure 4). Five birds responded more to the two prosodic cues (amplitude and duration) and three birds showed the opposite response. For none of the birds did this lead to a significant difference in response to the stimuli. This indicates that, for the current set of stimuli, pitch is the most salient cue to the birds.

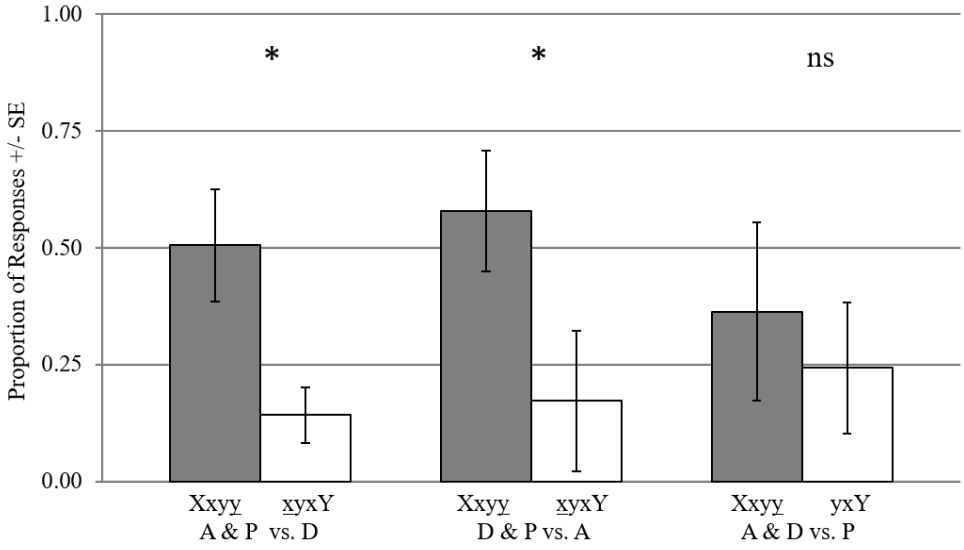


Figure 4. Proportions of responses to the test items in which two prosodic cues were not coinciding with the syntactic structure from the training. The remaining prosodic cue did correspond to the location of all prosodic features in the training. We show the mean proportion of response of all birds +s.e. The asterisks indicate a significant difference between the responses to the two stimuli types. Higher grey bars indicate a stronger response to the two prosodic cues. A, amplitude; D, duration; P, pitch.

Experiment 5: no prosody

Stimuli. As a final test, we used stimuli that had the exact same syllables and syntactic structure as during training, but were composed of unstressed syllables only. This allowed us to learn whether the birds did discriminate the syntactic structures as well.

Results and discussion. When only syntactic structure is available to discriminate the stimuli, the birds still discriminated the ones that followed the go structure from the ones that followed the no-go structure (mean responses xyxy = 0.17, mean responses xxyy = 0.09, $Z = 22.524$, $p = 0.012$). This indicates that although the birds are very responsive even to single prosodic cues, they also learned about the syntactic structure of the stimuli. It also demonstrates that a single prosodic cue can overrule the structural information.

GENERAL DISCUSSION

Our experiments show that, when presented with three prosodic cues, zebra finches are sensitive to the same prosodic features of human speech as humans are. Moreover, the zebra finches respond stronger to the prosodic features of our stimuli than

humans do. When required to either follow the prosodic cues or the syntactic structure, human participants are split equally in which of these they relied upon, whereas all of our zebra finches follow the prosodic pattern of the stimuli. The zebra finches are also able to generalize this prosodic pattern to quadruplets consisting of new syllables and to quadruplets that had a mismatch between the prosodic pattern, even with only a single prosodic cue present, and the syntactic structure. This is shown by an increased number of go-responses to stimuli in which the prosodic pattern matched with that of the go-stimuli of the training phase. Furthermore, the zebra finches do not differ in their responses to the stimuli in which only pitch, amplitude or duration is changed; they responded equally often to the stimuli in which the prosodic pattern followed the go-pattern of the training phase. When the cues are contradicting each other, pitch is the only cue that balances out the other two prosodic cues. However, it is notable in some experiments that the responses of the zebra finches to test items are in general lower than the responses to the training stimuli. This might have been owing to a novelty affect, causing the birds to respond in a risk aversive manner when confronted with new sounds.

Although animals can discriminate strings with different syntactic structures, the generalization of such structures to novel items has rarely been obtained in non-human animals [24,27]. Songbirds have only limited abilities to generalize different structures to novel sounds [24,27,28]. In contrast to this limited generalization based on structural cues, we show that zebra finches can readily generalize prosodic patterns to strings with novel syllables. This generalization suggests that the distinction learned during the training was not based on trivial cues related to the individual stimuli used for training, but on something more general that is shared by all stressed syllables, like the relative difference in pitch, amplitude or duration of the stressed syllable compared with the other ones. This ability for generalization or abstraction of the prosodic cues may also underlie the results obtained in a study on Java sparrows [16], where the sparrows were able to abstract the prosodic patterns of natural human speech sentences. Our results add to these findings by demonstrating that each of three different prosodic cues can be used to detect prosodic patterns. There was no detectable difference in the discrimination abilities of the birds when they heard a change in pitch, amplitude or duration as the only cue of the prosodic pattern.

When the zebra finches were presented with contradicting prosodic cues, we found a stronger effect of the pitch change compared with the other two prosodic features. In our experiment, pitch outweighed the other two prosodic cues. It is known that birds in general, and also zebra finches, are better at detecting small differences in frequencies than proportionally similar changes in duration or intensity in their natural

song [11,19,20]. Our results show that changes in pitch are also strong cues when songbirds listen to human speech syllables.

These results add to previous research showing that zebra finches are sensitive to phonetic cues by different speakers [21,22], by showing a sensitivity to other natural variations in human speech. They support the idea that such sensitivities either have a shared ancestral state or have evolved independently along highly similar lines in different groups of vocal communicators. The sensitivity of the birds to the features shown here could be related to the impact that these features might have in song interactions among conspecific birds. In the song of many songbirds, variations in these particular features can carry important information on, for example, the quality of the singing male [29]. When subjected to different environments or situations, songbirds are capable of modifying prosody-like features in their song [7,9]. For example, when food availability is decreased, an individual zebra finch often reduces its singing speed and song amplitude [30]. These song modifications show that at least some prosodic variation in birdsong is related to environmental or genetic parameters, indicating that these prosodic features could be of importance in vocal communication.

The subtle variation in vocal signals, which can be significant in communication, may also be the evolutionary origin of human sensitivity to such cues. One of the hypotheses concerning the evolution of human language is that it was preceded by a prosodic-protolanguage [31]. This protolanguage was based on a learned and generative vocal system with modifications in prosodic features as pitch, amplitude and duration. It might have its origins in a pre-existing sensitivity to meaningful variation in non-speech sounds present in our ancestors. The evolutionary process that shaped the sensitivity to relevant and sometimes subtle sound variations in birds, but most likely also that in other animals, might also have been at the base of the sensitivity to prosodic, paralinguistic features in human speech.

CONCLUSION

Results of our experiments showed that zebra finches are able to abstract and generalize prosodic patterns of human speech to new tokens. The zebra finches responded more strongly to the prosodic pattern than to the syntactic structure of the quadruplets. This response pattern persisted even when only one of three cues was present in a stimulus. Humans share this sensitivity to the prosodic pattern, although they also noted and responded to the syntactic structure of our stimuli. Our results

show that sensitivity to prosodic cues is not linked to the possession of language and might have preceded language evolution, possibly originating from a pre-existing sensitivity to meaningful variation in pre-linguistic communicative sounds.

Acknowledgements

We thank Andreea Geambasu, Raquel G. Alhama, Clara Levelt and Willem Zuidema for useful comments on the experimental procedure and on the manuscript. We thank Anouk de Weger, Stefan Wever and Dexter Wessels for help with the practical work.

Funding statement

This research was supported by NWO-GW, grant no. 360.70.452.

Ethical Standard

Research was approved by the Leiden Committee for animal experimentation DEC number 13007.

APPENDIX A

Training and test stimuli used for the experiments. A–H are the eight speech syllables that we used, these were different syllables for every individual. P, elevated pitch; D, increased duration and A, increased amplitude of the syllable.

Training Stimuli

Go-Stimuli	No-Go Stimuli
$a^{PDA}bab$	$aabb^{PDA}$
$c^{PDA}dcd$	$ccdd^{PDA}$
$e^{PDA}fef$	$eeff^{PDA}$

Test Stimuli

Experiment 1. Prosody versus Syntax

$a^{PDA}abb$	$abab^{PDA}$
$c^{PDA}cdd$	$cdcd^{PDA}$
$e^{PDA}eff$	$efef^{PDA}$

Experiment 2. Generalization

$h^{PDA}ghg$	$hhgg^{PDA}$
$g^{PDA}hgh$	$gghh^{PDA}$

Experiment 3. One Prosodic Cue

$a^{D}abb$	$abab^{D}$
$c^{D}cdd$	$cdcd^{D}$
$e^{D}eff$	$efef^{D}$
$a^{A}abb$	$abab^{A}$
$c^{A}cdd$	$cdcd^{A}$
$e^{A}eff$	$efef^{A}$
$a^{P}abb$	$abab^{P}$
$c^{P}cdd$	$cdcd^{P}$
$e^{P}eff$	$efef^{P}$

Experiment 4. Two Prosodic Cues versus One Prosodic Cue

$a^{AP}abb^{D}$	$a^{D}bab^{AP}$
$c^{AP}cdd^{D}$	$c^{D}dcd^{AP}$
$e^{AP}eff^{D}$	$e^{D}fef^{AP}$
$a^{PD}abb^{A}$	$a^{A}bab^{PD}$
$c^{PD}cdd^{A}$	$c^{A}dcd^{PD}$
$e^{PD}eff^{A}$	$e^{A}fef^{PD}$
$a^{DA}abb^{P}$	$a^{P}bab^{DA}$
$c^{DA}cdd^{P}$	$c^{P}dcd^{DA}$
$e^{DA}eff^{P}$	$e^{P}fef^{DA}$

Experiment 5. No Prosody

abab	aabb
cdcd	ccdd
efef	eeff

REFERENCES

1. Kuhl PK. 2004 Early language acquisition: cracking the speech code. Nat. Rev. Neurosci. 5, 831 – 843. (doi:10.1038/nrn1533)

2. Johnson EK, Seidl AH. 2009 At 11 months, prosody still outranks statistics. Dev. Sci. 12, 131 – 141. (doi:10.1111/j.1467-7687.2008.00740.x)

3. Jusczyk PW, Houston DM, Newsome M. 1999 The beginnings of word segmentation in English learning infants. Cogn. Psychol. 39, 159 – 207. (doi: 10.1006/cogp.1999.0716)

4. Seidl A, Johnson EK. 2006 Infant word segmentation revisited: edge alignment facilitates target extraction. Dev. Sci. 9, 565 – 573. (doi:10. 1111/j. 1467-7687.2006.00534.x)

5. Thiessen ED, Saffran JR. 2003 When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. Dev. Psychol. 39, 706 – 716. (doi:10.1037/0012-1649. 39.4.706)

6. Nazzi T, Jusczyk PW, Johnson EK. 2000 Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. J. Mem. Lang. 43, 1 – 19. (doi:10.1006/jmla.2000.2698)

7. Brumm H, Slater PJB. 2006 Animals can vary signal amplitude with receiver distance: evidence from zebra finch song. Anim. Behav. 72, 699–705. (doi: 10.1016/j.anbehav.2006.01.020)

8. Sossinka R, Bohner J. 1980 Song types in the zebra finch (Poephila guttata castanotis). Z. Tierpsychol. 53, 123 – 132. (doi:10.1111/j. 1439-0310.1980.tb01044.x)

9. Brumm H, Zollinger SA. 2011 The evolution of the Lombard effect: 100 years of psychoacoustic research. Behaviour 148, 1173 – 1198. (doi:10.1163/ 000579511X605759)

10. Brumm H, Ritschard M. 2011 Song amplitude affects territorial aggression of male receivers in chaffinches. Behav. Ecol. 22, 310–316. (doi:10. 1093/ beheco/arq205)

11. Nagel KI, McLendon HM, Doupe AJ. 2010 Differential influence of frequency, timing, and intensity cues in a complex acoustic categorization task. J. Neurophysiol. 104, 1426–1437. (doi:10. 1152/jn.00028.2010)

12. Ramus F, Hauser MD, Miller C, Morris D, Mehler J. 2000 Language discrimination by human newborns and by cotton-top tamarin monkeys. Science 288, 349–351. (doi:10.1126/science.288.5464.349)

13. Tincoff R, Hauser M, Tsao F, Spaepen G, Ramus F, Mehler J. 2005 The role of speech rhythm in language discrimination: further tests with a non-human primate. Dev. Sci. 8, 26-35. (doi:10.1111/j.1467-7687.2005.00390.x)

14. Toro JM, Trobalon JB, Sebastian-Galles N. 2003 The use of prosodic cues in language discrimination tasks by rats. Anim. Cogn. 6, 131 – 136. (doi:10. 1007/s10071-003-0172-0)

15. Toro JM,Trobalon JB, Sebastian-Galles N.2005 Effects of backward speech and speaker variability in language discrimination by rats. J. Exp. Psychol. Anim. B 31, 95 – 100. (doi:10.1037/0097-7403.31.1.95)

16. Naoi N, Watanabe S, Maekawa K, Hibiya J. 2012 Prosody discrimination by songbirds (Padda oryzivora). PLoS ONE 7, e47446. (doi:10.1371/ journal.pone.0047446)

17. Bolhuis JJ, Okanoya K, Scharff C. 2010 Twitter evolution: converging mechanisms in birdsong and human speech. Nat. Rev. Neurosci. 11, 747 – 759. (doi:10.1038/nrn2931)

18. Doupe AJ, Kuhl PK. 1999 Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. 22, 567 – 631. (doi:10.1146/annurev. neuro.22.1.567)

19. Lee TTY, Charrier I, Bloomfield LL, Weisman RG, Sturdy CB. 2006 Frequency-range discriminations and absolute pitch in black-capped chickadees (Poecile atricapillus), mountain chickadees (Poecile gambeli), and zebra finches (Taeniopygia guttata). J. Comp. Psychol. 120, 217–228. (doi: 10.1037/ 0735-7036.120.3.217)

20. Weisman RG, Balkwill LL, Hoeschele M, Moscicki MK, Bloomfield LL, Sturdy CB. 2010 Absolute pitch in boreal chickadees and humans: exceptions that test a phylogenetic rule. Learn. Motiv. 41, 156–173. (doi:10.1016/j.lmot. 2010.04.002)

21. Ohms VR, Gill A, Van Heijningen CAA, Beckers GJL, ten Cate C. 2010 Zebra finches exhibit speaker-independent phonetic perception of human speech. Proc. R. Soc. B 277, 1003 – 1009. (doi:10.1098/rspb. 2009.1788)

22. Ohms VR, Escudero P, Lammers K, ten Cate C. 2012 Zebra finches and Dutch adults exhibit the same cue weighting bias in vowel perception. Anim. Cogn. 15, 155–161. (doi:10.1007/s10071-011-0441-2)

23. Dooling RJ, Best CT, Brown SD. 1995 Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (Melopsittacus undulates) and zebra finches (Taeniopygia gutata). J. Acoust. Soc. Am. 97, 1839 – 1846. (doi:10.1121/1.412058)

24. van Heijningen CAA, de Visser J, Zuidema W, ten Cate C. 2009 Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. Proc. Natl Acad. Sci. USA 106, 20 538 – 20 543. (doi: 10.1073/pnas.0908113106)

25. Boersma P, Weenink D. 2014 PRAAT: doing phonetics by computer (computer program), v. 5.3.56. See http://www.praat.org/ (accessed 24 January 2013).

26. Koopmans van Beinum FJ. 1980 Vowel contrast reduction: an acoustic and perceptual study of Dutch vowels in various speech conditions. Amsterdam, The Netherlands: Academische Pers B.V.

27. ten Cate C, Okanoya K. 2012 Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. Phil. Trans. R. Soc. B 367, 1984–1994. (doi:10.1098/rstb. 2012.0055)

28. van Heijningen CAA, Chen J, van Laatum I, van der Hulst B, ten Cate C. 2013 Rule learning by zebra finches in an artificial grammar learning task: which rule? Anim. Cogn. 16, 165 – 175. (doi:10.1007/ s10071-012-0559-x)

29. Spencer KA, MacDougall-Shackleton SA. 2011 Indicators of development as sexually selected traits: the developmental stress hypothesis in context. Behav. Ecol. 22, 1 – 9. (doi:10.1093/beheco/ arq068)

30. Ritschard M, Brumm H. 2012 Zebra finch song reflects current food availability. Evol. Ecol. 26, 801–812. (doi:10.1007/s10682-011-9541-3)

31. Fitch T. 2005 The evolution of language: a comparative review. Biol. Philos. 20, 193–230. (doi:10.1007/s10539-005-5597-1)

Michelle Spierings, Carel ten Cate, Jeroen Hubert &
Henkjan Honing

Published in Frontiers in Psychology, 2016, 7:730.



5

CAN BIRDS PERCEIVE RHYTHMIC PATTERNS? A REVIEW AND EXPERIMENTS ON A SONGBIRD AND A PARROT SPECIES



ABSTRACT

While humans can easily entrain their behavior with the beat in music, this ability is rare among animals. Yet, comparative studies in non-human species are needed if we want to understand how and why this ability evolved. Entrainment requires two abilities: (1) recognizing the regularity in the auditory stimulus and (2) the ability to adjust the own motor output to the perceived pattern. It has been suggested that beat perception and entrainment are linked to the ability for vocal learning. The presence of some bird species showing beat induction, and also the existence of vocal learning as well as vocal non-learning bird taxa, make them relevant models for comparative research on rhythm perception and its link to vocal learning. Also, some bird vocalizations show strong regularity in rhythmic structure, suggesting that birds might perceive rhythmic structures. In this paper we review the available experimental evidence for the perception of regularity and rhythms by birds, like the ability to distinguish regular from irregular stimuli over tempo transformations and report data from new experiments. While some species show a limited ability to detect regularity, most evidence suggests that birds attend primarily to absolute and not relative timing of patterns and to local features of stimuli. We conclude that, apart from some large parrot species, there is limited evidence for beat and regularity perception among birds and that the link to vocal learning is unclear. We next report the new experiments in which zebra finches and budgerigars (both vocal learners) were first trained to distinguish a regular from an irregular pattern of beats and then tested on various tempo transformations of these stimuli. The results showed that both species reduced the discrimination after tempo transformations. This suggests that, as was found in earlier studies, they attended mainly to local temporal features of the stimuli, and not to their overall regularity. However, some individuals of both species showed an additional sensitivity to the more global pattern if some local features were left unchanged. Altogether our study indicates both between and within species variation, in which birds attend to a mixture of local and to global rhythmic features.

INTRODUCTION

In 1871, Darwin wrote: *"The perception, if not the enjoyment, of musical cadences and of rhythm is probably common to all animals and no doubt depends on the common physiological nature of their nervous systems"* (1). At the time, this thought was understandable as many animal species show behaviors that are characterized by some form of rhythmicity. It can be found in invertebrates, like the flashing patterns of fireflies, which can even be synchronized (2), as well as in vertebrates, like the strong rhythmicity characterizing some bird vocalizations. For instance, the cooing of the collared dove (*Streptopelia decaocto*) consists of a series of repeated "coos," each consisting of three vocal elements of different duration separated by brief pauses, also of unequal duration. This temporal pattern, and hence the structure of the coo as a unit, is highly stereotyped (3), resulting in a characteristic rhythmic pattern for a series of coos. Receivers are very sensitive to the overall regularity of the coo: if the temporal structure is changed, the responses are strongly reduced (4). The question is whether, as Darwin implied, such examples indicate a sensitivity to rhythmicity in general (ranging from a sensitivity for rhythmic pattern, pulse, and meter, as well as the influence of tempo and timing; 5) or whether this sensitivity is confined to particular species specific behaviors. Below, we will first review this topic, with particular attention to the "vocal learning and rhythmic entrainment hypothesis" formulated by Patel et al. (6-9). Doing so, we focus on studies on the perception of rhythmic patterns in birds, which for various reasons provide an ideal group for comparative studies on this topic. Next we present experimental data on pattern perception and the responsiveness to tempo changes in a songbird (zebra finch) and a parrot species (budgerigar).

The Vocal Learning and Rhythmic Entrainment Hypothesis

The interest in rhythm perception in animals is part of the more general quest for searching for signs of musicality in non-human animals, as a means to get more insight in the evolutionary and causal processes underlying human musicality (10, 11). The specific question whether animals can detect regularity in a stimulus and synchronize their own behavior to arbitrary rhythmic patterns got sudden attention with the discovery of Snowball, a sulfur-crested cockatoo that could synchronize head and body movements with the beat in several popular songs. Even though Snowball's behavior was only synchronized with the music for part of the time, he could adjust his movements to tempo changes of the songs (7-9). Parrots, such as Snowball, are vocal learners and vocal learning is associated with evolutionary modifications to the

forebrain, which plays a key role in mediating a link between auditory input and motor output during learning (12). As such linkage between auditory and motor areas in the brain is also required for beat induction (the ability to perceive a regular pulse in a varying rhythm, or real music; 5) and audio-motor entrainment, Patel et al. (8, 9) suggested that only vocal learning species might be able to show beat entrainment. A survey of YouTube movies searching for evidence of animal species that could entrain their behavior to music (7) seemed to confirm this hypothesis: entrainment was only observed among those species that showed vocal learning, suggesting that vocal learning was a necessary, albeit not a sufficient, requirement for beat induction. However, further studies have shown the picture to be more complicated. Convincing evidence of entrainment with a musical beat has now also been established in a Californian sea lion, named Ronan (13). Although, sea lions belong to a clade of mammals (pinnipeds) that contains some vocal learners (14), there is currently no evidence of vocal learning in this specific species, which potentially falsifies the generality of the hypothesis (15). There is also some evidence of chimpanzees adapting their finger tapping to an external beat, although this seems limited to frequencies close to their spontaneous motor tempo (16, 17). Chimpanzees are considered vocal non- learners, although it can be argued that they show some vocal plasticity and adjustment (18), and hence that their limited abilities to synchronize match with their limited abilities for vocal learning. In addition, recent evidence for temporally coordinated rhythmic movements between a bonobo (also a vocal non-learning species) and a human drummer (19) suggests that the link between vocal learning and beat induction may be less clear than initially anticipated. However, the most intriguing feature of the survey of Schachner et al. (7) is that of those taxa that show vocal learning (for mammals: dolphins and whales, seals, bats, and elephants; for birds: parrots, songbirds, and hummingbirds—12, 20), evidence for beat induction was only present in several parrot species and elephants. With respect to the latter, the evidence that elephants are vocal learners originates from captive Asian elephants, which imitated truck sounds (21), and words of the caretaker (22). However, in this latter study, the speech sounds were produced by inserting the trunk into the mouth, i.e., in a way quite different from how elephants usually produce vocalizations. As it is possible to teach elephants to perform behavior patterns well outside their natural range by operant conditioning, it might well be that the speech imitations also arose by operant shaping of the vocalizations by the human caretaker, hence being based on a different mechanism from the auditory imitative vocal learning in other species. Taken together, this leaves the parrots as the only group showing both imitative vocal learning and beat induction. This calls for a re-examination of the link between vocal learning and beat perception and induction.

Vocal Learning and Beat Perception Revisited: Are Parrots Special?

The YouTube survey (7) shows a remarkable contrast between the parrots and other vocal learning birds (hummingbirds and songbirds). Seven different parrot species all show evidence of beat induction. Another parrot species that is not listed as observed to synchronize with music, the budgerigar, has since been shown to be able to peck a key in synchrony with a flashing light and a metronome, and could learn to adjust this pecking to some tempo changes (23), although the adjustment to each new tempo was not spontaneous but was trained specifically. In striking contrast, the list of vocal learners contains 10 different songbird species and a hummingbird, none of which provided evidence for beat induction. In addition, three songbirds were erroneously classified under "vocal nonmimics" (nuthatch, bulbul, and babbler), with none of them showing beat induction. Thus, perhaps the question should be: why is it that various parrots, but no other vocal learning (or non- learning) birds, show beat induction? One possibility is that this difference is accidental. For instance, the total number of parrot movies is higher than that for the other bird species together, hence there may be a sampling bias. Or the difference might be related to behavioral differences between parrots and songbirds. Many parrot species show head bobbing or other body movements in their social interactions with conspecifics. If they are hand reared, as happens often with parrots, much of their social behavior will be directed to their human caretakers as a result of sexual or social imprinting (24) and a possible scenario might be that if their caretakers are dancing and moving on the beat, the parrots might be induced to do the same thing. Songbirds often lack such conspicuous rhythmic body movements in their natural behavior and may have less strong bonds with their caretakers as even captive ones are usually raised by their parents. Hence, they may possibly be less likely to provide evidence for beat detection and induction, even though they might be able to it. But it may also be that there is a more fundamental difference between parrots and other species. Showing beat induction requires at least two abilities: first the detection of a rhythmic pattern or beat in an external stimulus and next adjusting the frequency of some motor pattern to this input. Lack of beat induction may indicate lack of either or both of these abilities in other species. Thus, perhaps other bird species can detect rhythmic patterns in external stimuli, but lack the possibility to synchronize their behavior with it (see also 25).

Alternatively, non-parrot bird species might lack the ability for detecting rhythmic patterns in auditory stimuli altogether, due to the differences in brain pathways. It has been claimed that parrots have an enhanced vocal learning system, due to an extra song system that surrounds the song system shared with songbirds (26). The motor pathway system that surrounds this shell song system shows gene expression profiles

similar to the song system. These non-vocal motor brain regions are active during hopping and head bobbing movements (27) and it is therefore suggested that this motor system is involved in entrainment (26). Finally, the observed relationship between vocal learning and beat induction in parrots may be coincidental: parrots are vocal learners and show beat induction, but both may not be causally related or the relation may be due to some shared third factor underlying both.

To conclude: it is clear that it still is an open question why beat induction among birds has only been observed for parrots, calling for a further exploration of the topic. The observed contrast between parrots and songbirds make birds a particularly interesting group for comparative studies. Also, there are many vocal non-learning bird species, such as doves and pigeons, that show strong rhythmicity in their vocalizations and, finally, there are species not showing such rhythmicity. So comparing different species belonging to various avian groups may help to clarify the relation between vocal learning, beat and rhythm perception and beat induction. This may also reveal whether there is a categorical jump from synchronizing to pulses in natural behavior such as in flashing fireflies, to showing beat perception and synchronization, as implied by Patel et al. (8, 9).

Recently, Arriaga et al. (28) and Petkov and Jarvis (12) proposed a vocal learning continuum hypothesis to accommodate the different levels of vocal learning ranging from vocal non-learners (like doves), limited vocal learners to complex-vocal learners (like parrots and songbirds). It may well be that there is also a more fine grained spectrum in rhythmic patterns (see also 29), in particular when we shift the research focus from the production to the perception of rhythmic patterns. As the detection of rhythmic patterns, such as a pattern of repetitive identical inter-pulse-intervals or a higher order repetitive regularity in a rhythm, seems a first requirement for being able to move in synchrony with a beat, the central question in this paper concerns whether and which birds can detect such regularities in auditory patterns. And if so, is there a difference between species or groups of species in this ability? Or between vocal learners and non-learners?

Can Birds Detect Rhythmic Patterns and Regularity in Auditory Patterns?

In the first instance, asking whether birds, or any animal species, can detect regularity in an auditory pattern of, say, repeated pulses seems trivial. Studies on habituation have shown that various animals habituate more quickly to isochronous pulse series than to heterochronous ones (e.g., mice—30; zebra fish—31). However, this need not imply that they detect "isochrony" as such, as the distinction can be achieved by attending to local features, such as the differences in pause duration of one or a few

inter-pulse intervals and by predicting the timing of a next event from the preceding interval, i.e., is based on a sensitivity to absolute, and not relative, timing (32-34). Detecting isochrony as such, or rhythmic patterns more generally, involves a global process: detecting that events are regularly distributed over a longer series, irrespective of, for instance, the precise duration of the inter-pulse-intervals (see also 35). Rhythmic pattern detection thus concerns detecting a relational property. So, if an animal can distinguish between a regular, isochronous, pattern and an irregular, heterochronous one, a critical test to see whether this is based on having detected the global difference in regularity of the pattern is to see whether the discrimination is maintained after tempo transformations. So, what is known about such perceptual abilities in birds? To date, only a handful of bird species have been examined experimentally for their discrimination between isochronous and heterochronous sound patterns and/or for whether this discrimination is maintained with tempo changes of structured sound patterns. These are the domestic pigeon and several songbirds (starling, jackdaw, zebra finch). We briefly review these studies below.

Pigeons. Two studies examined whether pigeons could discriminate and generalize across different tempos. The first one (36) showed that pigeons subjected to a non-differential training in which they had to peck a response key to get food while being exposed to a regularly spaced train of pulses, generalized their response to slower, and faster pulse trains. This has also been demonstrated for quail chicks (37), but it tells little on the ability to detect regularity or even tempo generalization in general, as the birds may have attended to the mere presence of any sound. Differential training, in which responses to one pulse rate but not to another one, were rewarded resulted in discrimination between the two rates and the differentiation was maintained with stimuli showing either higher or lower pulse rates than the training ones. However, this study did not examine the ability to discriminate regular from irregular rhythms.

A more recent study examined, among others, whether pigeons are able to detect and discriminate different meters (38). Using two sounds, different meters were constructed using the same pulse rate (180bpm). The pigeons were able to discriminate the meters, but only if these differed substantially from each other (8/4 vs. 3/4). Further tests suggest that the pigeons might not have attended to the meter, but to the time difference between the beats. They did not transfer the discrimination to similarly structured stimuli consisting of other sounds, suggesting that their responses were also tied to the nature of the sounds. A second experiment on meter discrimination showed that the discrimination was maintained with faster tempos (200 and 220bpm), but not with reduced tempos of the pulse (140 and 160bpm). In a next experiment, the same birds were tested for their ability to discriminate an isochronous

from an irregular pulse pattern. The pigeons did not succeed in this discrimination. Finally, it was examined whether they could discriminate between two different isochronous pulse rates (with pulse and pause durations scaled proportionally), similar to the study by Farthing and Hearst (36). Three out of the four pigeons managed this discrimination and this was again generalized to slower and faster pulse trains. From these experiments, Hagmann and Cook (38) conclude that, on the whole, pigeons were most likely attending to the intervals between pulses, rather than to the overall metric or regular structure of the sound strings.

Starlings. Starlings (songbirds) were tested for the perception of regularity and rhythm in a series of studies by Hulse S. H. et al. (39), Hulse S. et al. (40), Humpal and Cynx (41). The birds were able to discriminate an isochronous (pulse duration 100 ms, intervals 100 ms) as well as a hierarchical pattern (four regularly spaced pulses followed by a longer pause and next followed by repetitions of this pattern) from a randomly generated heterochronous pattern with fluctuating pulse and pause durations. The discrimination was maintained with tempo changes in which pulse durations and intervals were extended or reduced proportionally (ranging from halving to doubling the tempo), although the strength of this discrimination was reduced for slower tempos (39, 40). The discrimination was reduced when the inter-pulse-interval remained constant, and pulse durations varied, but not the other way around (39, 40). The discrimination was also affected if pulse duration, but not the interval, was randomized or inverted, although it remained above chance level (41). Changing the pitch of the sounds affected the discrimination only slightly. Finally, the studies showed that starlings could discriminate two different rhythmic patterns consisting of four notes of different durations (50–50–300–300 ms vs. 50–300–300–50 ms), separated by longer pauses. Tempo transformation affected this discrimination, although it remained above chance in most cases (39, 40). These experiments suggest that starlings are better than pigeons in attending to more global patterns of pulse trains, although most experiments show some loss of discrimination with various tempo transformations.

Jackdaws. A pioneering study on rhythmic perception in jackdaws (corvid songbirds) was done by Reinert (42). He showed that a jackdaw could discriminate between two different auditory patterns with the structure ABAB and ABB respectively (A and B being different sounds). Among a series of other manipulations, he showed that the discrimination was maintained under tempo transformations (tempo training stimuli: 84bpm; test stimuli: 66–192bpm). A second jackdaw, trained to discriminate two other patterns, also maintained the discrimination with tempo transformations. Furthermore, the study demonstrated that the jackdaws maintained the discrimination

between the patterns when the sounds making up the patterns were changed (varying timbre or pitch), suggesting that the jackdaws used relative and global, rather than local features like specific interval durations or tone characteristics, to distinguish the patterns. However, the jackdaws have not been tested with isochronous stimuli in different tempos, nor whether they could discriminate an isochronous from an irregular pattern. Also, the stimuli used in this study were always very short strings. So, although suggestive, conclusive evidence that jackdaws are really sensitive to an overall rhythm formed by a repeated pattern is still lacking.

Zebra Finches. Zebra finches are a model songbird species for behavioral (e.g., 43-46) and neural studies on song learning (e.g., 47-49) as well as for comparative studies examining their abilities to discriminate various (speech) sounds or artificial grammar patterns (e.g., 50). Also, a few studies examined their abilities for detecting or discriminating rhythm- like structures. Nagel et al. (51) showed that zebra finches can distinguish the songs of two different males across various tempo transformations. Zebra finches can also detect prosodic patterns of edited speech sounds (52) and can discriminate song elements arranged in an ABAB structure from an AABB structure (53), and ABA structures from AAB structures (54, 55). Finally, exposure to a repeated series of regularly or irregularly spaced song elements induced differences in ZENK expression in two nuclei of the auditory system (NCM, CMM) (56). These observations suggest that zebra finches might also be able to discriminate between different rhythmic patterns or between regular and irregular pulsed sounds and maintain this with tempo transformations. However, although zebra finches can discriminate a regular isochronous from an irregular stimulus, this discrimination was strongly reduced with tempo transformations (changing the inter-pulse- intervals, but not the pulse durations), even if the training consisted of several tempo variants of the isochronous and irregular stimuli (57). These data suggested that the zebra finches, like pigeons, attended strongly to specific local features of the individual stimuli, such as the exact duration of inter-pulse intervals, rather than the overall regularity of the stimuli. Whether they are able to use more overall features still remains to be demonstrated (see 57, and below for a discussion).

To summarize the above overview: both pigeons and starlings are able to discriminate between two isochronous patterns in different tempos and maintain this discrimination with slower and faster tempos. However, this ability does not require perception of regularity as such, but can be achieved by attending to the duration of just one or a few intervals and generalizing from this to intervals that are more extreme to either one or the other end of the spectrum than the training ones. Although the currently available evidence is still limited, it suggests that all tested

species can solve discrimination tasks when this can be done by attending to such local temporal features of the sounds, suggesting this is the "default" state birds use for auditory pattern recognition. This is also how pigeons discriminate (some) different metric patterns, generated by alternation of two types of sounds and how zebra finches discriminate regular from irregular pulse patterns (57). Starlings also attend to the durations of pulses and intervals when discriminating between isochronous and randomly spaced sounds varying in duration. However, their ability to maintain the discrimination over at least some tempo changes suggests that they might also be sensitive to the larger pattern. This may also be true for the jackdaws.

It can be concluded that the evidence that birds can attend to some more global "regularity" or "rhythmicity" as such, is still very limited. However, whatever evidence there is suggests that this ability may differ between species. The studies of Snowball, as well as some data of a gray parrot and the YouTube survey (7) indicate that at least some parrot species have a quite well developed perceptual sensitivity for rhythm. The above review suggests the jackdaw as possible additional songbird candidate, but suggests also that this ability is poorer or even marginal in other songbird and non-songbird species. But, the experiments on all species are still equivocal on the issue, and more systematic comparative studies, focusing in particular on the discrimination between, and the responses to tempo transformations of regular vs. irregular stimuli are urgently required. Our experiments described below are meant to shed more light on such perceptual abilities.

Can Zebra Finches and Budgerigars Perceive Structural Regularity?

In our experiments we compare a songbird species, the zebra finch, with a parrot species, the budgerigar, for their abilities to discriminate a regular, hierarchically structured stimulus from an irregular one. We chose the budgerigar because the study of Hasegawa et al. (23) suggests that they are able to entrain their behavior to an audiovisual stimulus. Being a parrot species, we expect that they might also be able to attend to the more global features of temporal patterns in a perceptual discrimination task. For zebra finches, the current evidence for detecting pattern regularities is ambiguous: the study of van der Aa et al. (57) suggest they attended only to local temporal features, but in the song discrimination study by Nagel et al. (51) they were maintaining discrimination under tempo transformations. However, two songs changed in tempo might still be discriminable by other features that remained largely invariant after a tempo change, such as differences in the phonology of specific elements.

In the current study, we use training patterns that are similar in their (regular) inter-pulse-intervals, but differ in which pulses are accented, hence in their beat pattern. They thus show some hierarchical structure, providing the opportunity to examine how various local as well as more global temporal parameters affect the discrimination between the stimuli and whether this differs between the species. The birds are trained to discriminate between two hierarchical pulse strings, a regular one with one beat in each four pulses and an irregular string with the beats located at irregular positions. Subsequently, we test whether they generalize this discrimination to strings with modifications in the position and rate of the beat. This approach tests various hypotheses about which local and/or global features might be used when discriminating between regular and irregular strings.

MATERIALS AND METHODS

Subjects

Six male zebra finches and three female budgerigars were tested in this experiment. All zebra finches were between 120 and 321 days post hatching, the budgerigars were between 2 and 3 years old at the start of the experiment. The zebra finches were not subjected to previous experiments. The budgerigars had been used in a discrimination task with human speech and zebra finch sounds. Before the experiment, the animals were housed in group living facilities on a 13.5/10.5 L/D schedule and had food, water, and cuttlebone *ad libitum*. During the experiment, the L/D schedule was maintained, except for short dark periods as part of the experimental procedure. Water and cuttlebone were still *ad libitum*, the food availability was part of the experimental procedure and was monitored daily to ensure a sufficient level of food intake. The experiments were conducted in accordance to the animal experimentation guidelines of Leiden University. The protocol was approved by the Leiden committee for animal experiments, under DEC number 14071.

Apparatus

All experiments were conducted in an operant conditioning cage [zebra finches: 70(l) × 30(d) × 45(h) cm, budgerigars: 70(l) × 60(d) × 60(h) cm]. Each operant cage was in a separate sound attenuated chamber and was illuminated by a fluorescent tube that emitted a daylight spectrum on a 13.5 L: 10.5 D schedule. A speaker (Vifa 10BGS119/8) was located 1 m above the center of the cage. The sound level was set to 70 dB at the location of the bird at the start of a trial (in front of sensor 1). The cage walls were made from wire mesh except for a plywood back wall which

supported two pecking keys with LED lights. A food hatch was located in between these two keys, easily accessible to the birds. Pecking the left key (sensor 1) elicited a stimulus and illuminated the LED light of the key on the right (sensor 2). Depending on the sound, the bird had to peck sensor 2 or had to withhold its response. A correct pecking response resulted in access to food for 10 s. and an incorrect response led to 15 s. of darkness. Pecks during the sound presentation were not recorded as a response.

Experimental Design

Shaping. All birds started the experiment with a shaping procedure to get acquainted with the apparatus and the Go/No-go paradigm. This consisted of a 24-h acclimatization period with an opened food hatch, followed by a Go/No-go shaping procedure with one zebra finch song (Go sound) and one song element (No- go sound). Shaping lasted until the birds reached the standard discrimination ratio (response to Go sounds >75%, response to No-go sound <25%) for three consecutive days.

Discrimination Training. After shaping, all birds were trained to discriminate between one regular and one irregular string in the Go/No-go procedure (Figure 1). This training phase lasted until the bird reached the standard discrimination criterion for at least three consecutive days, after which it proceeded to the test phase.

Test Phase. During the test phase, test strings were randomly played at 20% of the trials, whilst the other 80% of the trials remained training strings test strings (see Figure 2). Feedback in the form of food access and darkness was only given during training trials, never for the test trials. All test strings were presented randomly and the test phase lasted until each test string was presented 40 times.

Stimuli

Stimuli were strings consisting of two different tones, an (accented) X-element (4000 Hz, 40 ms and 80 dB) and a y- element (2500 Hz, 40 ms, and 70 dB, created in Praat version 5.4.01), separated by a short silent interval, the pause (40 ms). The elements were concatenated to form two hierarchically organized training strings: one regular (the Go sound) and one irregular (the No-go sound, Figure 1), each lasting 3.5 s in total. For both strings the interval between the elements was identical, what differed between them was the position of the X-elements, which affected the number of y-elements between two X-elements. The regular string was a concatenation of 10 equal units, where each unit consisted of one X-element, followed by three y-elements

(Xyyy), all spaced equally. This concatenation created a stable 320 ms inter-X-interval (IXI), measured from onset to onset. The irregular string contained the same number of X- and y-elements as the regular string, but differed from this by variance in the number of y-elements between two X-elements. This variation ranged between one and five y-elements, creating a variation in the IXI between 160 and 480 ms. Both strings started with an additional three y-elements and had a fade-in and fade-out of 800 ms.

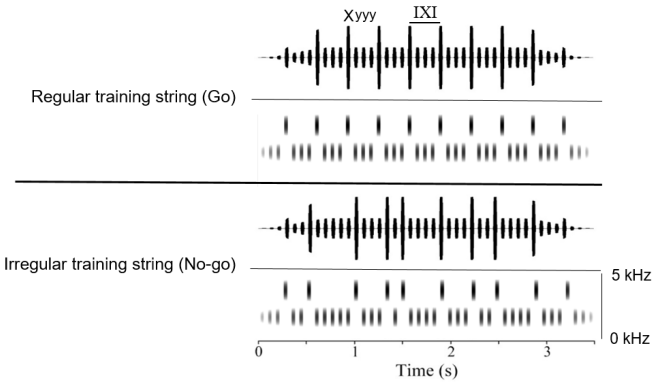


Figure 1. Visualization of the regular and the irregular training strings. The X-elements are indicated by the taller lines and the y-elements by the shorter lines. The IXI is the interval between two X-elements, measured from onset to onset. Note the fade-in and fade-out phases at beginning and end of the stimuli.

Test strings were created in a similar fashion as the training strings. They contained modifications in the duration and number of elements and pauses, thereby modifying the IXI and string length, while leaving the regular or irregular structure intact. Three main test sets were designed to systematically assess the effect of (1) modifications in the presence and location of elements, (2) the duration of elements and pauses, and (3) the number of elements on pattern detection (Figure 2, all regular strings used in this study are added as Supplementary Material).

Beat recognition: The role of number and presence of X- and y-elements

In these tests the IXI was kept identical to the training strings, whilst the number of y-elements in the string varied. If the birds discriminated the training strings by attending only to the duration of the accented pulses, i.e., the IXI interval, it is expected that varying the number of y-elements between the X's would not affect the discrimination. The test stimulus pair 1a (Figure 2.1) had an additional y-element between every two X-elements (four instead of three in the regular string). Test stimulus pairs 1b and 1c had a reduction of respectively one and three y-elements between two X-elements, creating regular strings with two and zero y-elements per IXI respectively. In the irregular strings each IXI is modified by adding or removing the same number of y-elements, with the limitation that there are never more than 5 or less than one y-element between two X-elements. Shortening or lengthening of the

pause durations compensated for these modifications and kept the IXI identical to the training strings. An additional test (pair 1d) was ran with only y-elements and prolonged pauses to compensate for the lack of X-elements (Figure 2.1).

Proportional scaling: The role of element and pause duration and IXI

In this test the number of X- and y-elements was identical to the training stimuli (Figure 2.2). However, the IXI interval was varied because the stimuli for this experiment had modified durations of the elements and pauses, both shorter and longer than the elements in the training stimuli. The regularity and irregularity of the training strings stayed intact by equally modifying all elements or all pauses in a string. If the birds are attending to the regularity of always having four y-elements between the X-elements in the regular string, the discrimination should be maintained. Reduced responding would indicate that the zebra finches attend to finer temporal details of the stimuli. Two versions of this modification were created.

For stimulus pair 2a both the elements as well as the pauses were lengthened with 25%. Pair 2b had the elements and pauses shortened by 25% (Figure 2.2). The strings of pair 2c had the pauses shortened with 50%, but the elements stayed identical to those in the training strings. For pair 2d the elements were shortened with 50%, whilst the pauses stayed identical to the training strings. This reduced the IXI of pair 2c and 2d with 25%, similar to test pairs 2b (Figure 2.2).

Numerical scaling: The role of the number of y-elements and IXI

In this test the IXI's were extended and compressed to the same length as in test 2, by adding (pair 3a) or removing (pair 3b) one y-element per X-interval (Figure 2.3). This manipulation created strings identical in the numbers of y-elements between X-elements to test stimuli 1a and 1b, but in this case the duration of the elements and pauses remained identical to those of the training stimuli, creating strings with a smaller or larger IXI (Figure 2.3). This stimulus thus maintains the finer details of element and pause durations from the training strings and only moves the location of the X-element within the string.

An assumption underlying the training and test procedures and stimuli is that humans exposed to these stimuli would recognize the regularity of the stimuli without being explicitly told to do so, and that, after training, they classify all test stimuli appropriately. To validate this assumption, we trained a group of 24 adult human participants to discriminate between the training strings and tested them with test set 1 and 2. The participants convincingly discriminated the regular from the irregular

strings of all test pairs (average response to regular stimulus = 0.88, average response to the irregular stimulus = 0.08, pairwise comparisons per test, all $p < 0.01$, see Supplementary Material). This indicates that, at least for humans, the regularity of the IXI intervals is recognizable, discriminable and generalizable. Thus far, we know that the sensitivity to temporal changes in birds, in the form of discriminating differences in duration or minimum integration time, is of a comparable level to that in humans (58, 59).

Analyses

The response data of the zebra finches and budgerigars was recorded as binomial measurements (number of Go and No-go responses). For the analysis, these measurements were converted to fractions between 0 and 1, calculated as the cumulative Go responses toward the Go or No-go strings, divided by the total number of trials. For the zebra finches, these fractions were analyzed with a generalized linear model (glm) with test item (all test strings and the training go and no-go string) as fixed effect and the individual as the random measure. This gave a significant effect of the test item on the Go-fraction ($t = 2.9$, $p = 0.004$). Pairwise comparisons were made between the fractions of responses to the Go and the No-go string of each test set and between the responses to the training and test strings by using a Tukey's *post-hoc* test, corrected for multiple testing. All results shown in the *Results Section* originate from these *post-hoc* tests. Furthermore, we ran a glm on all individual data to analyze the response pattern of each zebra finch by using the binominal response measures to each of the 40 trials per test string per bird. Results of this glm showed a significant effect of test item on the test scores ($t = 3.09$, $p < 0.001$) and results shown further are from the pairwise Tukey's *post-hoc* tests, restricted to only pairwise comparisons within each individual.

As only three budgerigars were tested, these data were only analyzed at the individual level. Like the zebra finch results, the responses of each budgerigar to each test string were measured as a binomial response. With a glm it was tested whether these scores differed over the test strings. The glm again showed a significant effect of test item on the test score ($t = 2.76$, $p = 0.005$). This was followed by a Tukey's *post-hoc* test with pairwise comparisons between the responses to the Go and No-go strings and between training and test strings within each individual. Results shown in the Result Section are from these *post-hoc* Tukey's test. All statistics were performed in Rstudio (version 0.98.1103).

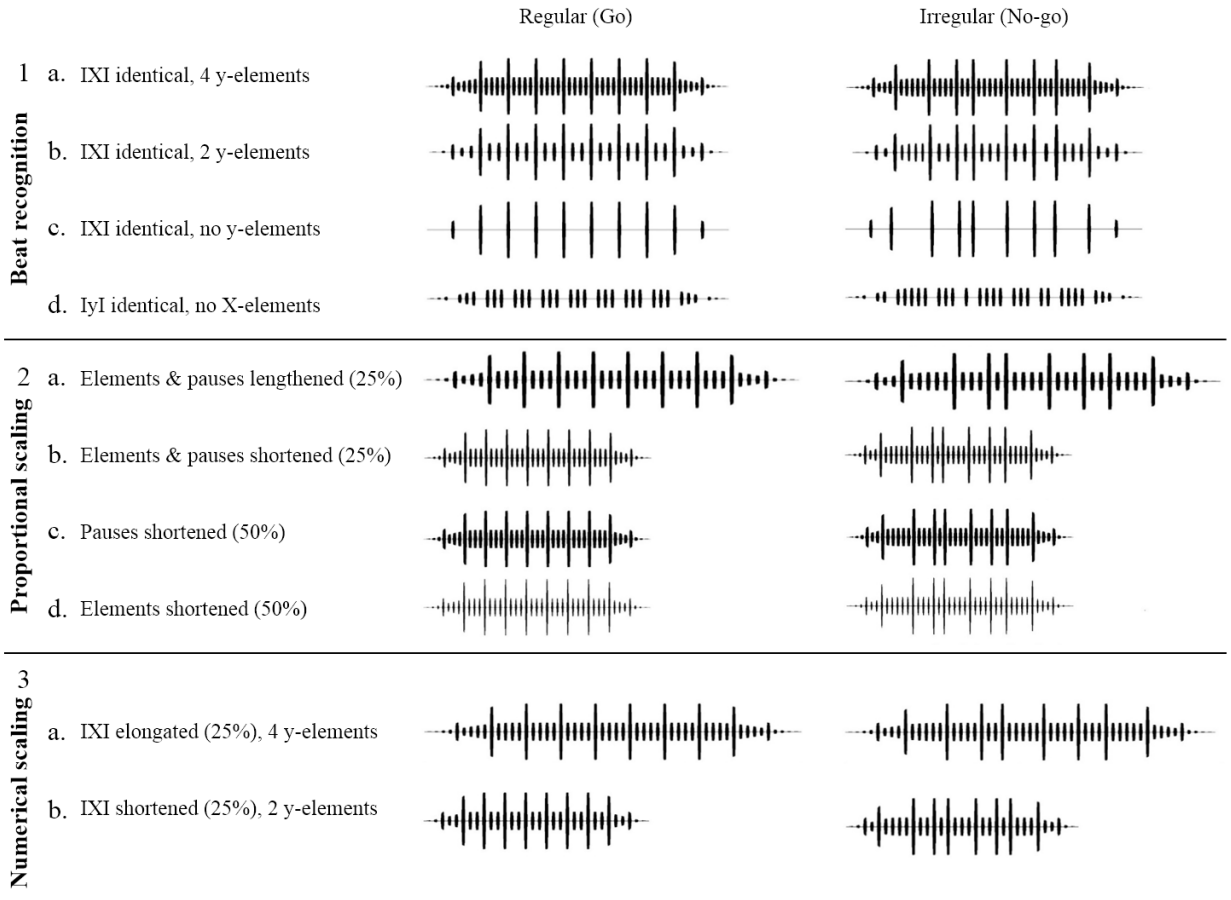


Figure 2. Visualizations of the different test pairs with the regular and the irregular string variants, grouped in three test sets. The longer vertical lines indicate the X-elements, shorter lines indicate the y-elements.

RESULTS

Training

The zebra finches required on average 10,245 trials to accurately discriminate between the regular and the irregular stimulus and complete the training. The three budgerigars learned the discrimination in 8495 trials on average.

Beat Recognition

Maintaining IXI but varying the number and presence of y- elements reduced the discrimination between the regular and irregular test stimuli (Figure 3). Zebra finches showed a trend toward a discrimination between a regular and an irregular string when these strings had one additional y-element within each IXI compared to the training strings (4 y-elements, pair 1a: $z = -5.08$, $p = 0.08$). A discrimination bordering significance is shown when the strings have one y-element less in each IXI (2 y-elements, pair 1b) with elongated pauses ($z = -3.37$, $p = 0.05$).

Strings consisting of only X-elements with an identical IXI as the training strings (pair 1c) resulted in a reduction in the responses to the regular string, which was no longer discriminated from the irregular string ($z = 0.02$, $p = 0.81$). A similarly low number of responses and no discrimination was recorded when only the y-elements of the training strings were present in the test strings (pair 1d: $z = -1.56$, $p = 0.22$).

Two zebra finches (Z2 and Z5) correctly discriminated the strings with one y-element more in each IXI (pair 1a, Z2: $z = -3.98$, $p < 0.01$; Z5: $z = -4.65$, $p < 0.01$), whilst the other four zebra finches did not discriminate (all $z > -2.13$, $p > 0.25$). None of the individuals discriminated between the regular and irregular string of pair 1b (one y-element less per IXI), 1c (no y-elements), and 1d (no X-elements; all $z > -1.33$, $p > 0.5$).

The budgerigars also showed reduced responses to test stimuli with an identical IXI to training, but with modified numbers of y- elements (Figure 3). Nevertheless, one budgerigar (B1) correctly discriminated the regular and irregular strings both with 4 y-elements (pair 1a) and 2 y-elements (pair 1b) between the X-elements (pair 1a: $z = -7.23$, $p < 0.01$; pair 1b: $z = -6.14$, $p < 0.01$). The other two budgerigars (B3 & B2) did not discriminate these strings (pair 1a and pair 1b: both $z > 0.8$, $p > 0.9$).

Similarly to the zebra finches, budgerigars did not discriminate between the regular and irregular string when only the X-elements (pair 1c) or only the y-elements (pair 1d) were present (pair 1c: all $z > 0.3$, $p > 0.9$; pair 1d: all $z > 0.04$, $p > 0.9$).

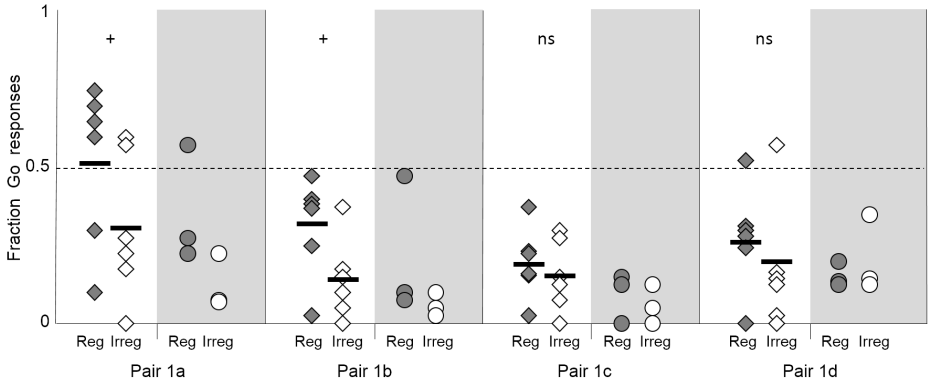


Figure 3. Fraction of Go-responses toward the regular test strings (shaded diamonds) and the irregular test strings (open diamonds) for the four pairs of test set 1. Horizontal bold lines shown the average fraction of Go responses of the six zebra finches. + symbols indicate a trend ($0.05 < p < 0.10$) toward a difference between the responses to the Go and to the No-go strings, ns indicates no significant difference (data for zebra finches only). Individual budgerigar results are shown with shaded circles (regular strings) and open circles (irregular strings). They were not tested at group level.

These results show that discrimination between regular and irregular strings was only partially maintained when the IXI remained constant whilst the number of y-elements varied. Only two zebra finches and one budgerigar discriminated between strings with one extra y-element in the IXI. None of the birds maintained the discrimination when only the y-elements and their intervals were present. It is clear that both element types were required and that whatever the birds might have used to discriminate the training strings, it was not just regularity, nor exact duration of the IXI.

Proportional Scaling

Modifications in the duration of both pauses and elements evoked different effects depending on the direction of the modification (Figure 4). Zebra finches showed no discrimination between the regular and irregular strings when both elements and pauses were elongated by 25% (pair 2a: $z = -1.73$, $p = 0.14$). However, they did make a correct discrimination between the regular and irregular string when both elements and pauses were shortened by 25% (pair 2b: $z = -6.31$, $p = 0.03$).

Keeping the element duration identical to training stimuli but shortening the pauses in the strings by 50% (pair 2c), showed a bordering significant result toward good discrimination by the zebra finches ($z = -3.98$, $p = 0.05$). A similar trend in the responses was found when the elements were shortened, but pauses kept similar to the training stimuli (pair 2d: $z = 3.29$, $p = 0.06$).

Two zebra finches (Z2 and Z5) discriminated the regular and irregular string with shorter pauses and elements (pair 2b, Z2: $z = -6.86$, $p < 0.01$; Z5: $z = -6.92$, $p < 0.01$, Figure 4). These were the same individuals that discriminated the regular and irregular string from pair 1a (4 y-elements per IXI). One zebra finch made a correct discrimination when only the pauses were shortened (pair 2c, Z5: $z = -5.47$, $p < 0.01$). Two other zebra finches correctly discriminated when the elements were shortened (pair 2d, Z2 and Z3: $z = -4.98$, $z = -4.09$, both $p < 0.01$).

Budgerigars hardly responded to strings with modified pause and element durations (Figure 4). Irrespective of the type of modification and whether the elements, the pauses, or both were modified, none of the budgerigars discriminated between the regular and the irregular strings (pair 2a: all $z > 0.03$, $p > 0.9$; pair 2b: all $z > 0.07$, $p > 0.6$; pair 2c: all $z > -1.4$, $p > 0.2$; pair 2d: all $z > -2.8$, $p > 0.4$).

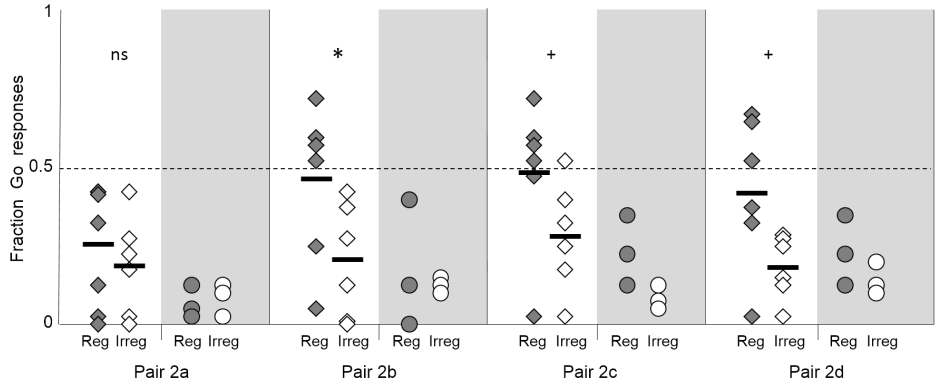


Figure 4. Fraction of Go-responses toward the regular strings (shaded diamonds) and irregular strings (open diamonds) with elongated (2a) or shortened (2b) elements and pauses, as well as only shortened pauses (2c) and only shortened elements (2d) creating a 25% increase or decrease of the IXI. Black lines shown the average fraction of Go responses of the six zebra finches. Asterisks indicate a significant difference between the responses to the Go and to the No-go strings, + symbols indicate a trend, ns indicates no significant difference. Individual budgerigar results are shown with shaded circles (regular strings) and open circles (irregular strings).

The duration of the elements and pauses influenced the birds' discrimination abilities differently in the two species. While budgerigars failed to discriminate between proportionally scaled strings, zebra finches' discrimination was maintained with shortened elements and pauses, although it was lost for all individuals when elements and pauses were elongated. There was no clear indication that reductions of elements, of pauses, or of both differ in their effect. The results suggest that the zebra finches showed at least some generalization of the discrimination when the number of y-elements between X-elements is left intact.

Numerical Scaling

The zebra finches maintained their discrimination between regular and irregular pulse strings when each IXI contained 4 y-elements, one y-element more than in the training strings, and thus had a 25% increase in duration of the IXI compared to training (pair 3a, Figure 5). Overall, zebra finches showed more Go-responses to the regular string than to the irregular string ($z = -9.88$, $p < 0.001$). The same discrimination ability was found when there were 2 y-elements in each IXI, creating a decrease in duration of the IXI by 25% (pair 3b: $z = -6.61$, $p = 0.002$). The level of discrimination between the regular and irregular string did not differ between these two manipulations ($z = 0.06$, $p = 0.45$).

All but one zebra finch discriminated between the strings with 2 y-elements and a shorter IXI (pair 3b, Z1, Z2, Z3, Z4, and Z5). Also, two zebra finches made this discrimination when there were 4 y-elements and the IXI was elongated (pair 3a, Z1: $z = -6.23$, $p < 0.01$, and Z4: $z = -8.45$, $p < 0.01$).

One budgerigar (B1) discriminated correctly when each IXI contained an extra y-element, creating an IXI increase of 25% ($z = -5.45$, $p < 0.01$), whilst the other two budgerigars did not discriminate these strings (both $z > -2.24$, $p > 0.7$, Figure 5). When the IXI was reduced by 25% by removing 1 y-element between the X-elements, again one budgerigar (B2) made a correct discrimination ($z = -4.33$, $p < 0.01$), while the other two budgerigars did not discriminate (both $z > 2.17$, $p > 0.15$).

These results confirm that the IXI did not need to be identical to the training strings for the birds to correctly discriminate between a regular and an irregular pulse string. In these test strings, the durations of the elements and pauses were maintained, but the number of y-elements varied. This also demonstrates that in this test the birds did not use the exact number of y-elements between two X-elements, nor the location of the X-element to discriminate between the training strings. Rather it seems that

generalization to longer and shorter regular patterns was at its best if the element and pause durations were kept identical to the training stimuli.

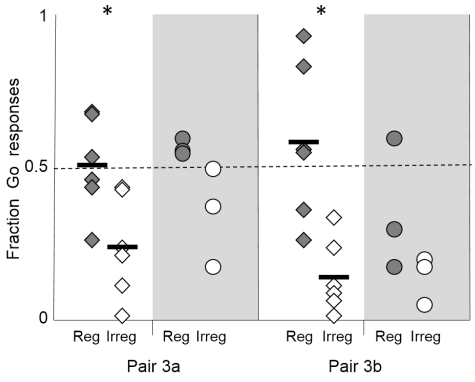


Figure 5. Fraction of Go-responses toward the regular strings (shaded diamonds) and irregular strings (open diamonds) with one y-element more (4 y-elements, pair 3a) or less (2 y-elements, pair 3b) compared to the training strings, creating a 25% increase or decrease of the IXI. Black lines shown the average fraction of Go responses of the six zebra finches. Asterisks indicate a significant difference between the responses to the Go and to the No-go strings. Individual budgerigar results are shown with shaded circles (regular strings) and open circles (irregular strings).

Comparing responses to training and test strings

A comparison between the responses of the zebra finches to the training and to the test strings revealed that although there were differences in the responses toward regular and irregular strings in the various tests, the average fraction of Go responses to the regular test strings was always lower than the responses to the regular training strings (pairwise comparisons regular test strings ~ regular training string, all $z < -9.92$, $p < 0.01$). Nevertheless, the zebra finches always responded more often to the regular test strings than they did to the irregular training strings (pairwise comparison regular test strings ~ irregular training, all $z < -4.54$, $p < 0.01$). The irregular test string of pair 1a (increased number of y-elements, identical IXI), pair 2c (pauses shortened by 50%), and pair 3a (IXI elongated by extra y-element) were the only stimuli to which the birds responded more often with a Go response than they did to the irregular training string (pairwise comparisons irregular test strings ~ irregular training string, pair 1a, pair 2c and pair 3a: $z < -5.68$, $p < 0.01$, all other $z > -1.7$, $p > 0.1$).

The budgerigars also responded less to all regular and irregular test strings than they did to the regular training strings (all $z < -5.78$, $p < 0.01$). However, some regular test strings got more Go responses than the irregular training strings. When one y-element was added between two X-elements and the IXI was increased correspondingly (pair 3a), all budgies responded with more Go responses to the regular string than to the irregular training string (all $z < -6.38$, $p < 0.01$). Additionally, one budgerigar (B2) also responded more strongly to regular test strings than to irregular training strings when there was one y-element removed (pair 3b), when there

were no y-elements (1c) or when both pauses and elements were elongated (2a) (all $z < -3.73$, $p < 0.02$). Budgerigar B3 responded stronger to the regular test string than to the No-Go training string when one y-element was added or removed, but the IXI stayed identical to training (pairs 1a and 1b, $z = -5.78$, $z = -5.21$, both $p < 0.01$).

DISCUSSION

Zebra finches and budgerigars can learn to discriminate between regular and irregular pulse strings in a Go/No-go operant training procedure. If the birds, like humans (see Supplementary Material), would make the discrimination based on differentiating on the basis of presence or absence of regularity, one would expect that all regular test stimuli would obtain similar Go-scores to the regular training stimulus, and be preferred consistently over the irregular test stimuli. This was not the case. Responding was considerably lower to regular test stimuli than to regular training stimuli and there is no consistent preference for the regular over the irregular test stimulus. However, several regular test stimuli got more responses than their irregular counterpart. So, what might underlie the differential responding?

Our three test-sets (see Figure 2) provide insights into the features of the regular and irregular training strings that zebra finches and budgerigars used when discriminating between them. The first test set showed that the birds did not discriminate the regular and irregular strings by attending exclusively to the IXI, nor by attending to the pattern of the y- elements. Apparently both element types are required to make the discrimination. However, some individuals maintained the discrimination with an increased number of y-elements and constant IXI. Test set two revealed that the zebra finches, but not the budgerigars, tended to maintain discrimination between the regular and irregular strings if the number of y- elements remained constant, but duration of pauses and/or elements were shortened. Discrimination was absent in both species when both elements and pauses were longer than the training strings. Finally, the third set showed that both zebra finches and budgerigars can discriminate between regular and irregular strings in which the number of y-elements and IXI is varied, provided that the duration of elements and pauses is maintained.

Concentrating on the statistically significant findings of the different individuals shows the presence of three main patterns. (1) Memorization without generalization: one budgerigar (B3) and one zebra finch (Z6) did not discriminate between any of the test string sets, suggesting that they memorized the training strings providing a food reward and discarded all deviating strings. (2) Generalization across varying IXI when

local features of the test strings, like element and pause length, were identical to the training strings. One budgerigar (B2) and three zebra finches (Z1, Z3, and Z4) discriminated strings with more or fewer y-elements between the X-elements and therefore a longer or shorter IXI (pair 3a and 3b). (3) Generalization with local variation: One budgerigar (B1) and two zebra finches (Z2 and Z5) discriminated strings with longer or shorter elements and IXI's, indicating that they were able to generalize regularity beyond local features. However, each individual had a specific subset of test strings which it discriminated, showing that there are still some specific local features that played a role during discrimination.

The individual variation among zebra finches, ranging from a focus on the exact structure of the stimuli to one with additional attending to a more global structure, has also been found in experiments in which zebra finches had to distinguish among string sets based on different artificial grammar patterns (53-55) and may hence reflect a variation in more general cognitive abilities. Our current results are also in line with the suggestion arising from reviewing earlier studies (see Section Introduction) that birds have a primary strategy to pay attention to local temporal features, in this case the duration of the elements and the pauses between them, for auditory pattern recognition. However, also in the present study it is clear that this initial strategy might be accompanied by a sensitivity to more global features, like the regularity of the pulse strings, as is shown by the correct discrimination between strings in which the IXI is modified by adding y-elements, but keeping identical element and pause durations (see pair 3a in Figure 5). Some sensitivity to regularity is also suggested by the finding that zebra finches responded more to the regular test strings than to the irregular training strings. The differentiation among the test stimuli of each type also suggests that they most likely based their responses on comparing test strings with both the regular as well as the irregular training string. In our experiment we used only a single regular and a single irregular training string. While this was sufficient for humans to classify novel strings as being regular or irregular this was not the case for the birds. However, it may be that if the birds had been trained on a set of regular and irregular stimuli they might have shifted more clearly from using local features to using the global feature of regularity.

Our zebra finch results seem somewhat in between those obtained by Nagel et al. (51) and those of van der Aa et al. (57) for the discrimination between two stimuli of which the temporal parameters were varied compared to the training stimuli. The study by Nagel et al. (51), using songs from two different males, showed that discrimination of manipulated stimuli was similar to those of the training stimuli with changes in song duration of even >25% (57) used one type of pulse, separated by

isochronous or heterochronous intervals and showed that discrimination between regular and irregular test stimuli disappeared with a 25% tempo change. Our stimuli were more complex than those of van der Aa et al. (57) by using two types of elements, but lacked the phonological features present in full songs. In the present study, a 25% tempo change did affect some, but not all of the discriminations. Zebra finch songs differ in many features, such as the pitch contours, element length, amplitude modulations and formant patterns. Some of these features might have remained recognizable in the study of Nagel et al. (51) where the songs were proportionally scaled, allowing the zebra finches to use these features, instead of the rhythmic ones. Hence we cannot be sure that the rhythmic structure of the songs was used in maintaining the discrimination in that study. The results of van der Aa et al. suggested that zebra finches attend in particular to local features, in that case the exact duration of inter-onset intervals. Our current results support this partly, as discrimination seems most affected when durations of pauses and elements were manipulated, but also show that some discrimination was maintained with a shortening, but not with a lengthening of element and pause durations. Maintenance of some discrimination between regular and irregular stimuli with proportional scaling of both pauses and elements has also been shown for starlings (39, 40) and pigeons (36). It is of interest that for both of these species a decrease in tempo resulted in a stronger reduction of discrimination than an increase, similar to what is observed in the current experiment. The starlings appeared more sensitive to changes in tone length than changes in inter-onset interval, while the zebra finches in our study seemed to give equal weight to both.

The reduction in discrimination resulting from proportional scaling was, for both zebra finches and budgerigars, stronger than that for starlings, which maintained good discrimination with a 40% tempo change (40). Hulse S. et al. (40) interpreted their findings as indicating at least some sensitivity to rhythmicity for starlings. Our results are less conclusive on this issue. They suggest that both zebra finches and budgerigars showed some sensitivity to stimulus regularity, but only when some local features remained invariant. Similar ambiguous findings were observed in other studies of rhythm perception in birds, as discussed in the introduction. For example, pigeons could discriminate between meters with different pulse rates and between different regular pulse strings, but not between a regular and an irregular pulse string (38). Furthermore, they could not generalize the meter discrimination to pulse strings with similar rhythmic features, but different sound items. These results indicate that the discrimination by the pigeons was based on local phonological and temporal features, such as the absolute inter-pulse-intervals, and not on the global regularity of the strings. In a follow-up experiment using some of the same birds as used in the current

experiment (Spierings et al., unpublished), we also found that both species hardly responded when X- and y-elements were replaced by elements of the same duration but differing in phonetic structure. In contrast to the pigeons, starlings (and possibly jackdaws; 42, see Section Introduction) were able to discriminate between regular and random pulse strings, and could generalize this discrimination to some modifications of these strings, indicating that they might have attended to the global rhythmic feature of the pulse strings (39-41). However, just like the zebra finches, the starlings also discriminated best if some local features remained identical to the training strings, such as pulse duration, whilst changes in others such as the pitch of the pulses, did not affect the discrimination.

Abilities related to auditory-motor rhythm production have so far mainly been shown in avian species belonging to the parrot clade. Not only larger parrots can, to a certain extent, synchronize their body movements with a beat, but also the smaller budgerigars have shown rhythmic entrainment (23). Nevertheless, this particular experiment might not have required regularity perception from the budgerigars. They were required to peck on a key at certain regular interval, indicated by a light and a sound. Repeating the previously heard or seen interval, i.e., attending to absolute timing, might have allowed the birds to correctly entrain to the presented rhythm. The budgerigars in the present study did not use the general regularity of the strings to discriminate, since they only discriminated between specific regular and irregular strings.

So, both zebra finches and budgerigars were in general not using the global regularity when discriminating these strings, but both could attend to some aspects of regularity. This is in contrast to another study, in which budgerigars and zebra finches were tested on their rule learning strategies (Spierings and ten Cate, in revision). That study showed that zebra finches used local, positional information to discriminate song element triplets (XYX and XXY), whilst budgerigars used a global strategy and attended to the structure of the strings. This resulted in a generalization of the structural rules by the budgerigars, but not by the zebra finches. One noticeable difference between that study and the current on is that Spierings and ten Cate (in revision) used a set of exemplars of the XYX and the XXY string during training, whereas in the current study the animals were trained with one regular and one irregular string. Less variation in training strings might have reduced the attention given to the general regularity-irregularity difference. Nevertheless, if the difference in regularity of the strings was the most prominent strategy to discriminate them, this strategy should also be employed with only one exemplar of each, as shown by the human subjects (see Supplementary Material).

One way of interpreting the existing literature and the current results is to distinguish between at least three types of perceptual biases that might characterize time and rhythm perception in birds and other animals. These three types are a bias for local features of auditory elements (such as pitch, amplitude, duration), a bias for more global prosodic features (such as pitch contour or amplitude contour), or a bias for the temporal structure, such as inter-beat-intervals. In the current study, most individuals seem to use local temporal features as their primary strategy in solving the discrimination task. We refer to this as the local feature bias hypothesis. This hypothesis suggests a preference in birds for local elements (such as duration, inter-onset interval, pitch, amplitude, or timbre) in perception and discrimination tasks and a lower sensitivity to whether they are part of a more global temporal structure, be it isochronous, heterochronous or metrical. This is not to say that zebra finches and budgerigars cannot take advantage of the global structure; it is just not their preferred strategy in solving this type of discrimination tasks.

To summarize the results of the current experiment and those reviewed in the introduction of our study: there is between and within species variation in how well different birds are able to detect regularity of pulse strings. However, while the vocal non- learning pigeons seem to perform poorest on this, there is only a gradual difference with vocal learners such as zebra finches and budgerigars, which in turn show a gradual difference with starlings and jackdaws. Also, if there is, as our review suggested, a difference between parrots and other bird species in sensitivity to regularity and rhythm, it does not hold for the budgerigar. Also, the currently available data show no systematic differences among vocal learners and non-learners. So, we suggest, similar to what Merchant and Honing (33) suggested for primates, that the current data show a continuum (instead of a categorical jump) in the ability to detect regularity and rhythmicity. This idea is similar to the continuum hypothesis suggested for vocal learning by Arriaga et al. (28) and Petkov and Jarvis (12). However, it should be realized that the number of species tested for their abilities to perceive regularity or rhythm is still limited and the test methods and stimuli varied. Hence, there is a need to extend experiments to other avian groups, both vocal non- learners as well as some vocal learning groups that are considered to be more advanced in their cognitive abilities (such as large parrots and corvids) and therefore may be expected to have more elaborate rhythm perception.

Acknowledgements

We thank Guus van der Velden and Sissy Bijsterbosch for help with the animal work. Furthermore, we thank the two referees for their constructive comments. This research was supported by NWO-GW, grant 360.70.452.

SUPPLEMENTARY INFORMATION

To verify that humans treat our stimuli by attending to their regularity, adult human participants were trained and tested in a Go/no-go paradigm on a computer. The participants were instructed on the functioning of the go/no-go, and were subsequently trained and tested with the same strings as were used for the birds.

METHODS

Subjects

Twenty-four human participants (age 29.5 +/- 2.7) were trained and tested in this experiment, one extra participant was unable to successfully finish the training. All participants were native Dutch speakers, reported to have no hearing problems.

Apparatus and experimental design

The participants were asked to sit in front of a computer and received a set of headphones (Denon DJ DN-HP1000) with a sound level of 70dB. The Go/No-go paradigm was written in Processing, a script based design program. Before the start of the experiment participants were instructed to click at a speaker icon that appeared in the middle of the screen. This click started a trial and the playback of a stimulus, followed by the appearance of a blue rectangle. The participants were instructed to either click or not click on this rectangle before it would disappear after 3.5 seconds. Feedback was provided with a happy smiley and a "ping" sound after a correct Go response and a sad smiley and a loud low sound after an incorrect Go response.

After the instruction, the participant started the training. Participants had to learn to discriminate between a regular string (Go sound) and an irregular string (No-go sound, figure 1). After they reached the standard discrimination ratio, more than 75% correct Go responses and less than 25% incorrect No-Go responses within a block of 30 trials, they continued to the test phase.

Participants received a written instruction on the screen notifying that they continued to the test phase and would not receive any feedback. Test items were presented in blocks of 20 trials, 16 test strings and 2 repetition of the 2 training strings. The test was finished when a participant finished 4 blocks of 20 trials.

Stimuli

The training stimuli were identical to those used for the zebra finches and budgerigars (figure 1). Also the same test stimuli were used, although humans were only tested on test set 1 and 2 (figure 2.1 and 2.2). This means they received test strings with one more y-element between two X-elements, but with an identical IXI (inter-X-interval) to the training (pair 1a). Strings with one y-element less between two X intervals, with an identical IXI (pair 1b). Strings with no y-elements between two X-interval, again the IXI stayed identical to training (pair 1c). And strings without X-elements, where the y-elements remained in identical positions compared to the training string (pair 1d).

In test set 2 the IXI were longer or shorter than the training strings. The participants heard strings with longer elements and pauses, creating an IXI increase of 25% (pair 2a) and strings with shorter elements and pauses, resulting in a decrease in the IXI of 25% (pair 2b). Furthermore, they heard a strings with only the pauses shortened, creating an IXI decrease of 25% (pair 2c), and a string with only the elements shortened to create an IXI decrease of 25% (pair 2d).

Analyses

The Go and No-go responses of the participants were calculated as fractions correct and fraction incorrect by the following formula's: Fraction correct = correct Go/ (correct Go + incorrect Go), fraction incorrect = incorrect Go/ (correct Go + incorrect Go). These fractions were then analyzed with generalized linear model and Tukey's post-hoc test in which the correct and incorrect fractions per test were pitted against each other.

RESULTS

24 out of 25 participants learned to discriminate between the regular and the irregular training string, they needed on average 1.7 blocks of 30 trials. One participant did not learn the discrimination within 120 training trials and did not proceed to the test.

Regardless of the test pair, humans discriminated between the regular and the irregular strings (figure S1, all $p<0.01$). However, the discrimination between the regular and the irregular string without X-elements (pair 1d), was less strong compared to the other tests (all comparisons with pair 1d $p<0.05$). No difference in discrimination strength was found between the other test pairs (all $p>0.1$).

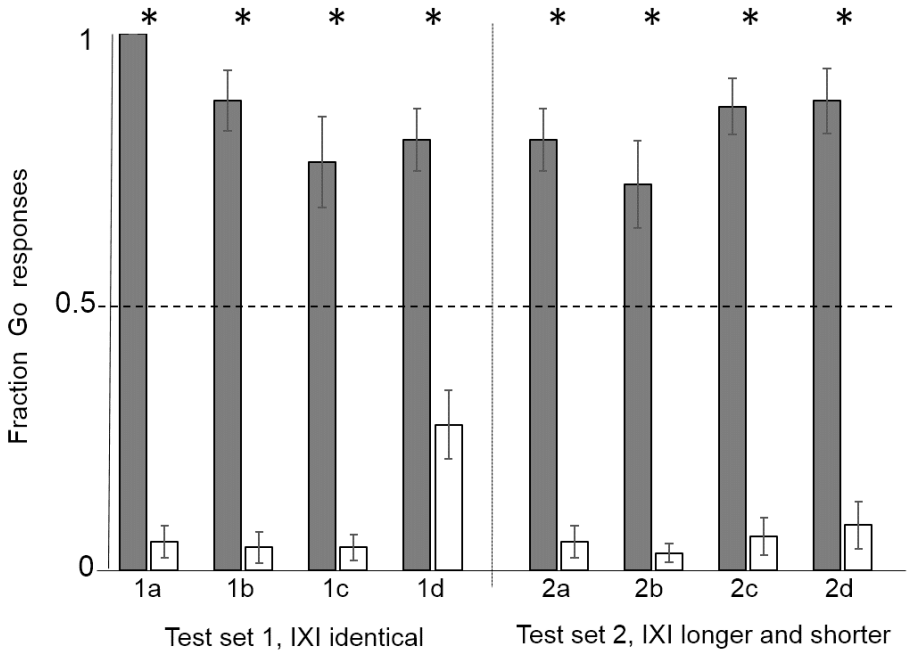


Figure S1. Responses of human participants to the different test pairs. The shaded bar shows the fraction of correct responses to regular (Go) strings, the open bar shows the fraction of incorrect Go responses to irregular (No-go) strings. Bars show the average score, whiskers the standard error of the mean. Asterisks indicate a significant difference obtained in Tukey's post-hoc tests from a GLM.

REFERENCES

1. Darwin, C. (1871). The Descent of Man, and Selection in Relation to Sex. London: John Murray.

2. Buck, J. (1988). Synchronous rhythmic flashing of fireflies. II. Q. Rev. Biol. 63, 265–289. doi: 10.1086/415929

3. Ballintijn, M. R., and ten Cate, C. (1999). Variation in number of elements in the perch-coo vocalization of the collared dove (Streptopelia decaocto) and what it may tell about the sender. Behaviour 136, 847–864. doi: 10.1163/156853999501603

4. Slabbekoorn, H., and ten Cate, C. (1999). Collared dove responses to playback: slaves to the rhythm. Ethology 105, 377–391. doi: 10.1046/j.1439-0310.1999.00420.x

5. Honing, H. (2013). "Structure and interpretation of rhythm in music," in The Psychology of Music, ed D. Deutsch (London: Elsevier/Academic Press), 369–404.

6. Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. Music Percept. 24, 99–104. doi: 10.1525/mp.2006.24.1.99

7. Schachner, A., Brady, T. F., Pepperberg, I. M., and Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. Curr. Biol. 19, 831–836. doi: 10.1016/j.cub.2009.03.061

8. Patel, A. D., Iversen, J. R., and Bregman, M. R. (2009a). Experimental evidence for synchronization to a musical beat in a nonhuman animal. Curr. Biol. 19, 827–830. doi: 10.1016/j.cub.2009.03.038

9. Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009b). Studying synchronization to a musical beat in nonhuman animals. Ann. N.Y. Acad. Sci. 1169, 459–469. doi: 10.1111/j.1749-6632.2009.04581.x

10. Hoeschele, M., Merchant, H., Kikuchi, Y., Hattori, Y., and ten Cate, C. (2015). Searching for the origins of musicality across species. Philos. Trans. B Biol. Sci. 370:20140094. doi: 10.1098/rstb.2014.0094

11. Honing, H., ten Cate, C., Peretz, I., and Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. Philos. Trans. B Biol. Sci. 370, 1–8. doi: 10.1098/rstb.2014.0088

12. Petkov, C. I., and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. Front. Evol. Neurosci. 4:12. doi: 10.3389/fnevo.2012.00012

13. Cook, P., Rouse, A., Wilson, M., and Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. J. Comp. Psychol. 127, 412–427. doi: 10.1037/a0032345

14. Reichmuth, C., and Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. Curr. Opin. Neurobiol. 28, 66–71. doi: 10.1016/j.conb.2014. 06.011

15. Wilson, M., and Cook, P. (2016). Rhythmic entrainment: why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. Psych. Bull. Rev. doi: 10.3758/s13423-016-1013-x. [Epub ahead of print].

16. Hattori, Y., Tomonaga, M., and Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. Sci. Rep. 3:1566. doi: 10.1038/srep01566

17. Hattori, Y., Tomonaga, M., and Matsuzawa, T. (2015). Distractor effect of auditory rhythms on self-paced tapping in chimpanzees and humans. PLOS ONE 10:e0130682. doi: 10.1371/journal.pone.0130682

18. Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., et al. (2015). Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees. Curr. Biol. 25, 495–499. doi: 10.1016/j.cub.2014. 12.032

19. Large, E. W., and Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a spontaneous tempo and rhythmic entrainment in a bonobo. J. Comp. Psychol. 129, 317–328. doi: 10.1037/com0000011

20. Janik, V. M., and Slater, P. J. B. (1997). Vocal learning in mammals. Adv. Study Behav. 26, 59–99. doi: 10.1016/S0065-3454(08)60377-0

21. Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., and Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. Nature 434, 455–456. doi: 10.1038/434455a

22. Stoeger, A. S., Mietchen, D., Oh, S., de Silva, S., and Herbst, C. T., Kwon, S., et al. (2012). An Asian elephant imitates human speech. Curr. Biol. 22, 2144–2148. doi: 10.1016/j.cub.2012.09.022

23. Hasegawa, A., Okanoya, K., Hasegawa, T., and Seki, Y. (2011). Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. Sci. Rep. 1:120. doi: 10.1038/srep00120

24. ten Cate, C., and Vos, D. R. (1999). Sexual imprinting and evolutionary processes in birds: a reassessment. Adv. Study Behav. 28, 1–31. doi: 10.1016/S0065- 3454(08)60214-4

25. Patel, A. D., and Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Front. Syst. Neurosci. 8:57. doi: 10.3389/fnsys.2014.00057

26. Chakraborty, M., Walløe, S., Nedergaard, S., Fridel, E. E., Dabelsteen, T., Pakkenberg, B., et al. (2015). Core and shell song systems unique to the parrot brain. PLoS ONE 10:e0118496. doi: 10.1371/journal.pone.0118496

27. Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., et al. (2008). Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. PLoS ONE 3:e1768. doi: 10.1371/journal.pone.0001768

28. Arriaga, G., Zhou, E. P., and Jarvis, E. D. (2012). Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song learning birds. PLoS ONE 7:e46610. doi: 10.1371/journal.pone.0046610

29. Ravignani, A., Martins, M., and Fitch, W. T. (2014). Vocal learning, prosody, and basal ganglia: don't underestimate their complexity. Behav. Brain Sci. 37, 570–571. doi: 10.1017/S0140525X13004184

30. Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K., et al. (2007). Processing of temporal unpredictability in human and animal amygdala. J. Neurosci. 27, 5958–5966. doi: 10.1523/JNEUROSCI.5218-06.2007

31. Shafiei Sabet, S., Neo, Y. Y., and Slabbekoorn, H. (2015). The effect of temporal variation in sound exposure on swimming and foraging behaviour of captive zebrafish. Anim. Behav. 107, 49–60. doi: 10.1016/j.anbehav.2015.05.022

32. Honing, H., and Merchant, H. (2014). Differences in auditory timing between human and nonhuman primates. Behav. Brain Sci. 37, 557–558. doi: 10.1017/S0140525X13004056

33. Merchant, H., and Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. Front. Neurosci. 7:274. doi: 10.3389/fnins.2013.00274

34. Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., and Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. Philos. Trans. B Biol. Sci. 370:20140093. doi: 10.1098/rstb.2014.0093

35. Geiser, E., Walker, K. M. M., and Bendor, D. (2014). Global timing: a conceptual framework to investigate the neural basis of rhythm perception in humans and non-human species. Front. Psychol. 5:159. doi: 10.3389/fpsyg.2014.00159

36. Farthing, G. W., and Hearst, E. (1972). Stimulus generalization and discrimination along the click-frequency (flutter) continuum in pigeons. Percept. Psychophys. 12, 176–182. doi: 10.3758/BF03212866

37. Schneider, S. M., and Lickliter, R. (2009). Operant generalization of auditory tempo in quail neonates. Psychon. Bull. Rev. 16, 145–149. doi: 10.3758/PBR.16.1.145

38. Hagmann, C. E., and Cook, R. G. (2010). Testing meter, rhythm, and tempo discriminations in pigeons. Behav. Process. 85, 99–110. doi: 10.1016/j.beproc.2010.06.015

39. Hulse, S. H., Humpal, J., and Cynx, J. (1984). Processing of rhythmic sound structures by birds. Ann. N.Y. Acad. Sci. 423, 407–419. doi: 10.1111/j.1749-6632.1984.tb23449.x

40. Hulse, S., Humpal, J., and Cynx, J. (1984). Discrimination and generalization of rhythmic and arrhythmic sound patterns by European starlings (Sturnus vulgaris). Music Percept. 1, 442–464. doi: 10.2307/40285272

41. Humpal, J., and Cynx, J. A. (1984). Discrimination of temporal components of acoustic patterns by birds. Ann. N.Y. Acad. Sci. 423, 600–602. doi: 10.1111/j.1749-6632.1984.tb23466.x

42. Reinert, J. (1965). Takt- und rhythmusunterscheidung bei dohlen. Z. Tierpsychol. 22, 623–671. doi: 10.1111/j.1439-0310.1965.tb01683.x

43. Slater, P. J. B., Eales, L. A., and Clayton, N. S. (1988). Song learning in zebra finches (Taeniopygia guttata): progress and prospects. Adv. Study Behav. 18, 1–34. doi: 10.1016/S0065-3454(08)60308-3

44. Jones, A. E., ten Cate, C., and Slater, P. J. B. (1996). Early experience and plasticity of song in adult male zebra finches (Taeniopygia guttata). J. Comp. Psychol. 110, 354–369. doi: 10.1037/0735-7036.110.4.354

45. Tchernichovski, O., Mitra, P. P., Lints, T., and Nottebohm, F. (2001). Dynamics of the vocal imitation process: how a zebra finch learns its song. Science 291, 2564–2569. doi: 10.1126/science.1058522

46. Lipkind, D., Marcus, G. F., Bemis, D. K., Sasahara, K., Jacoby, N., Takahasi, M., et al. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. Nature 498, 104–108. doi: 10.1038/nature12173

47. Jarvis, E. D., and Nottebohm, F. (1997). Motor-driven gene expression. Proc. Natl. Acad. Sci. U.S.A. 94, 4097–4102. doi: 10.1073/pnas.94.8.4097

48. Haesler, S., Wada, K., Nshdejan, A., Morrisey, E. E., Lints, T., Jarvis, E. D., et al. (2004). FoxP2 expression in avian vocal learners and non-learners. J. Neurosci. 24, 3164–3175. doi: 10.1523/JNEUROSCI.4369-03.2004

49. Zeigler, P., and Marler, P. (2008). The Neuroscience of Birdsong. Cambridge: Cambridge University Press.

50. ten Cate, C. (2014). On the phonetic and syntactic processing abilities of birds: from songs to speech and artificial grammars. Curr. Opin. Neurobiol. 28, 157–164. doi: 10.1016/j.conb.2014.07.019

51. Nagel, K. I., McLendon, H. M., and Doupe, A. J. (2010). Differential influence of frequency, timing, and intensity cues in a complex acoustic categorization task. J. Neurophysiol. 104, 1426–1437. doi: 10.1152/jn.00028.2010

52. Spierings, M. J., and ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. Proc. Biol. Sci. 281:20140480. doi: 10.1098/rspb. 2014.0480

53. van Heijningen, C. A. A., de Visser, J., Zuidema, W., and ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. Proc. Natl. Acad. Sci. U.S.A. 106, 20538–20543. doi: 10.1073/pnas.0908113106

54. van Heijningen, C. A. A., Chen, J., van Laatum, I., van der Hulst, B., and ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? Anim. Cogn. 16, 165–175. doi: 10.1007/s10071-012-0559-x

55. Chen, J., van Rossum, D., and ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. Anim. Cogn. 18, 151–164. doi: 10.1007/s10071-014-0786-4

56. Lampen, J., Jones, K., McAuley, J. D., Chang, S.-E., and Wade, J. (2014). Arrhythmic song exposure increases ZENK expression in auditory cortical areas and nucleus taeniae of the adult zebra Finch. PLoS ONE 9:e108841. doi: 10.1371/journal.pone.0108841

57. van der Aa, J., Honing, H., and ten Cate, C. (2015). The perception of regularity in an isochronous stimulus in zebra finches (Taeniopygia guttata) and humans. Behav. Processes eds D. Deutsch (London, UK: Elsevier/Academic Press) 115, 37–45. doi: 10.1016/j.beproc.2015.02.018

58. Dooling, R. J., Lohr, B., and Dent, M. L. (2000). "Hearing in birds and reptiles," in Comparative Hearing: Birds and Reptiles, eds R. J. Dooling, A. N. Popper, and R. R. Fay (New York, NY: Springer-Verlag), 308–359.

59. Dooling, R. J. (2004). "Audition: can birds hear everything they sing?," in Nature's Music, eds P. Marler and H. Slabbekoorn (San Diego, CA: Academic Press/Elsevier), 206–225.

Michelle Spierings & Carel ten Cate

Published in Frontiers in Neuroscience, 2016, 10:345.



6

ZEBRA FINCHES AS A MODEL SPECIES TO UNDERSTAND THE ROOTS OF RHYTHM

ABSTRACT

In this opinion piece we relate the rhythm perception abilities of zebra finches to two recently published papers on rhythms in the vocal production of zebra finches. Several experiments have shown that zebra finches can learn to discriminate between regular and irregular patterns. Recently, studies examining the vocal production of zebra finches showed that songs and call interactions often have a fixed, regular pattern. This relation of rhythmicity in perception and production shows that zebra finches might be an ideal model species fto study the relation between rhythm perception and production. I have added this opinion piece to this thesis, because it relates well to the studies on (rhythmic) pattern perception in zebra finches and budgerigars.

Vocal timing in zebra finches

Zebra finches are a widely used model species for neurobehavioural research, in particular in relation to song development and auditory processing. Males learn their songs from a tutor. Females don't sing, but do develop learned song preferences. Regardless of the differences, both sexes exchange calls in social interactions. Two fascinating recent studies looked at different aspects of rhythmicity in the production of zebra finch vocalizations (1, 2), and together with several studies on the perception of rhythms, they make the zebra finch a promising model species to unravel the roots of rhythm production and perception.

Norton & Scharff (2) analysed the intervals from one note onset to the next one in both isolate and directed zebra finch songs. For each male they were able to derive an isochronous sequence of 'time stamps' of which a subset aligned with all note onsets of a male's song. Moreover, these time stamps often also coincided with the transitions between phonetically different gestures within a complex note. This indicates that an isochronous rhythm might underlie zebra finch songs. Benichov and colleagues (1, see also 3) showed that both males and females can dynamically adjust the mutual timing of their calls. Individual birds were housed with a robotic zebra finch that emitted an isochronous call pattern. Within ample minutes the zebra finches adjusted their call rate to create a regular back-and-forth exchange with the robotic finch. The robot was then set to emit "jamming" calls that were produced at the moment when the zebra finch was most likely to respond. All zebra finches adjusted their call pattern to avoid the jamming calls, either by calling earlier, later or both earlier and later. Furthermore, when the robotic finch produced a pattern of alternating single and paired calls, the zebra finches timed their calls differently for the single compared to the paired calls, indicating they apparently detected the alternating pattern of the robot calls and used this to anticipate whether the next call would be single or paired. Interestingly, females performed better than males at these tasks (1).

Benichov et al. (1) next examined the role of the forebrain song system in the timing of calls. The female song system is reduced compared to the male system and lacks, for instance, the largest song nucleus, Area X. However, some nuclei of the song system are present in both sexes, like the RA, a nucleus related to the temporal aspects of zebra finch song learning. Lesioning the RA nucleus in zebra finches made them unable to correctly adjust the timing of their calls, although they were still responsive to the robotic finch. This indicates that for both males and females the RA is actively involved in call synchronization and vocal coordination. Further experiments showed that input to the RA from another forebrain nucleus, the HVC, was necessary to maintain call synchronization (1). While Norton and Scharff (2) did not examine the

neural basis of the isochronous patterns they observed in male songs, they also suggest that it might originate from the activation pattern of HVC neurons firing to the RA. These HVC_{RA} neurons fire in a rhythmic, clock-like pattern (4). However, as the rhythms observed by Norton & Scharff are three to ten times slower than the firing of these HVC neurons, they propose that additional mechanisms must operate to translate the clocklike firing into the higher complexity observed in the songs.

Whether animals can perceive beat and rhythmic patterns is a prominent question in relation to understanding the evolution of human musicality (5-9) and Benichov et al. (1, 3) suggest that the link between the neural system involved in vocal learning and call synchronization provides the relation between beat perception and auditory-motor coordination as shown in a number of animal species. So, how do the above findings on the presence of rhythms in vocal production in zebra finches relate to studies that addressed the perception of rhythmic auditory patterns?

Recognizing regularity

Rhythm can be defined as regular repeated pattern. In its simplest form it is an isochronous series of pulses, like the one produced by the robotic finch. More complicated rhythms can be created by repeating a heterochronous pattern, for example the repetition of single and double pulses as used in Benichov et al. (1). Humans are skilled in perceiving various rhythms. We can easily detect whether a pulse or beat pattern is regular or irregular by integrating temporal information over a series of auditory events (e.g. 10). We thus have the cognitive ability to abstract a general, global, pattern from a string of sounds, enabling us to classify patterns as regular or not. However, the question whether or to what extent non-human animals are also able to integrate temporal information over a longer series of sounds to classify strings as being regular is still open. A telling example is a study by Hagmann & Cook (11) that showed that pigeons were unable to learn to discriminate between an isochronous and an irregular pulse pattern. Pigeons are vocal non-learners and this might be the reason for their inability, as it has been suggested that vocal learning and vocal non-learning species might differ in this respect (12, 13; see also ten Cate et al. (14) for a review on this relationship for birds). The rhythm perception by zebra finches in the study by Benichov et al. seems to support this view, as they show that the (vocal learning) zebra finches extracted the regularity of the call pattern and used this to time their calling. However, when we combine these findings with several other recent studies on zebra finch perception of regularity the picture becomes more complex.

Nagel and colleagues (15) trained female zebra finches to discriminate between two different songs and showed that the females maintained the discrimination over a range of tempo changes. The songs were still categorized correctly up to a 25% speed increase or decrease. These results might indicate rhythm generalization by zebra finches. However, Nagel et al.'s conclusion was that zebra finches maintained the discrimination by attending to the spectral envelope of the songs. Attending to the sequence of local spectral features, rather than any timing pattern may have enabled the discrimination. Whether zebra finches do attend to regularity in the timing of songs was examined by Lampen et al. (16). They compared ZENK expression in response to playback of rhythmic zebra finch songs, i.e. where all songs in a string had identical inter-element intervals, with the expression in response to arrhythmic songs, i.e. a string of songs in which inter-element intervals vary. Arrhythmic songs resulted in stronger ZENK expression in the caudomedial nidopallium (NCM), the caudomedial mesopallium (CMM), and the nucleus taeniae (Tn). This increased activity in auditory areas of the zebra finch brain might be related to the finding by Benichov et al. (1) as, similar to their study, the repeated pattern in the regular song may have initiated predictive timing of the next song rendition, which would be lacking with the arrhythmic song.

Regularity or interval detection?

The above mentioned studies may indicate that zebra finches perceive 'regularity' as such. However, as also noticed by Benichov et al. (3), this need not be the case. The various findings may arise because repeated events with a fixed interval may give rise to a prediction for a next interval of the same absolute duration. So, when the birds respond to the robotic finch they can do so by attending to, and learning about, the absolute interval between successive robot calls, or by detecting other local contiguities of events (3), without having formed some concept of 'regularity'. The same accounts for the study by Lampen et al. (16), in which the zebra finches could also have responded to identical consecutive intervals in a sound string.

If zebra finches can perceive regularity as such, one would expect that they are not only able to distinguish a regular from an irregular string, but also to transfer this distinction to strings with modified tempos. This has been tested in experiments in which van der Aa et al. (17) trained zebra finches to discriminate between a set of regular, isochronous pulse strings and a set of irregular pulse strings. The irregularity in these strings was created by varying the duration of inter-pulse-intervals within a string. As expected from the studies discussed above, zebra finches of both sexes could learn to discriminate these strings. But this discrimination broke down when the zebra finches were tested with probe strings with novel tempo transformations. The

birds seemed to distinguish and discriminate the different strings based on their specific inter-pulse-interval durations, suggesting that the zebra finches focused on local features, without attending to or learning about the global pattern of regularity-irregularity of the strings (17).

Another recent experiment on zebra finches used isochronous pulse strings with pulses of two types. These were present in a ratio of 1:4, with the rare type raised in both frequency and amplitude compared to the other (14). This variation was used to create a string with a fixed number of low tones between two high ones, creating a regular beat pattern, and a string with different numbers of low tones in each interval, making the beat pattern irregular. Again, the zebra finches discriminated these strings. They were next tested with strings in which the location of the beat within the strings or the duration of pulses and inter-pulse-intervals was changed. This revealed that here also, the zebra finches seemed to use different local features, like the inter-beat- or inter-pulse-intervals to distinguish the strings. However, some birds seemed to combine a sensitivity to such local features with one for the more global regularity (14). Nevertheless, the behavioural experiments so far do not indicate that zebra finches perceive the global pattern of regularity as such. We should therefore be cautious with concluding that the ability of zebra finches (and other non-human species) to distinguish regular from irregular sound patterns is similar to the ability of humans to detect rhythm. However, there may be a continuum among species ranging from those not being able to discriminate regular from irregular sounds up those able of beat detection in more complex rhythms (14).

Conclusion and outlook

The various studies discussed above show that zebra finches are very good at detecting fixed interval durations and can use preceding intervals to predict the next one in a string of sounds. This matches their ability to call and to produce songs with a fixed rhythmic periodicity. Furthermore, specific nuclei of the zebra finch forebrain song system (HVC, RA, CMM and NCM) seem involved in producing and detecting rhythmicity. However, the results also call for further behavioural and neural studies on the links between perception and production of rhythmic patterns in the zebra finch. Are receivers sensitive to the rhythms underlying songs? And which brain areas are involved in the discrimination of the rhythmic patterns in the experiments of van der Aa et al. (17) and ten Cate et al. (14)? Are these perceptual abilities and also the production of the song rhythms observed by Norton & Scharff (2) affected by interfering with HVC and RA? The zebra finch has proven to be an excellent model species, and is very suitable for addressing these questions. However, we also need comparative studies on other species to understand the types of temporal patterns

that birds can detect, how they do so and how this is related to vocal learning (see also 3). Ultimately, this may also shed light on the building blocks from which our human ability for rhythm perception and beat entrainment may have evolved.

REFERENCES

1. Benichov, J.I., Benezra, S.E., Vallentin, D., Globerson, E., Long, M.A., and Tchernichovski O. (2016a). The forebrain song system mediates predictive call timing in female and male zebra finches. Curr. Biol. 26, 309–318.

2. Norton, P., and Scharff, C. (2016). 'Bird song metronomics': Isochronous organization of zebra finch song rhythm. Front. Neurosci. 10:309.

3. Benichov, J.L., Globerson, E., and Tchernichovski, O. (2016b). Finding the beat: From socially coordinated vocalizations in songbirds to rhythmic entrainment in humans. Front. Hum. Neurosci. 10:255.

4. Long, M.A. and Fee, M.S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature, 456, 189-194.

5. Patel, A. D. (2006). Musical Rhythm, Linguistic Rhythm, and Human Evolution. Music Percept. An Interdiscip. J. 24, 99–104.

6. Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. Front. Syst. Neurosci. 7, 68.

7. Patel, A. D. (2014). The evolutionary biology of musical rhythm: was Darwin wrong? PLoS Biol. 12, e1001821.

8. Honing, H., ten Cate, C., Peretz, I., and Trehub, S.E. (2015). Without it no music: Cognition, biology, and evolution of musicality. Phil. Trans. R. Soc. B 370: 20140088.

9. Hoeschele, M., Merchant, H., Kikuchi, Y., Hattori, Y., and ten Cate, C. (2015). Searching for the origins of musicality across species. Phil. Trans. R. Soc. B 370: 20140094.

10. Geiser, E., Walker, K.M.M., and Bendor, D. (2014). Global timing: a conceptual framework to investigate the neural basis of rhythm perception in humans and non-human species. Front. Psychol. 5(159), 1-7.

11. Hagmann, C.E., and Cook, R.G. (2010). Testing meter, rhythm, and tempo discrimination in pigeons. Behav. Process. 85(2), 99-110.

12. Schachner, A., Brady, T. F., Pepperberg, I. M., and Hauser, M. D. (2009). Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. Curr. Biol. 19, 831–836.

13. Patel, A. D., Iversen, J. R., and Bregman, M. R. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. Curr. Biol. 19, 827–830.

14. ten Cate, C., Spierings, M.J., Huberts, J., and Honing, H. (2016). Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species. Front. Psychol. 7:730.

15. Nagel, K. I., McLendon, H. M., and Doupe, A. J. (2010). Differential influence of frequency, timing, and intensity cues in a complex acoustic categorization task. J. Neurophysiol. 104, 1426–1437.

16. Lampen, J., Jones, K., McAuley, J. D., Chang, S.-E., and Wade, J. (2014). Arrhythmic song exposure increases ZENK expression in auditory cortical areas and nucleus taeniae of the adult zebra Finch. PLoS One 9, e108841.

17. Van der Aa, J., Honing, H., and ten Cate, C. (2015). The perception of regularity in an isochronous stimulus in zebra finches (Taeniopygia guttata) and humans. Behav. Process. 115, 37-45.

Michelle Spierings & Carel ten Cate

Published in PNAS, 2016, 113(27), E3977–E3984.



7

BUDGERIGARS AND ZEBRA FINCHES DIFFER IN HOW THEY GENERALIZE IN AN ARTIFICIAL GRAMMAR LEARNING EXPERIMENT



ABSTRACT

The ability to abstract a regularity that underlies strings of sounds is a core mechanism of the language faculty but might not be specific to language learning or even to humans. It is unclear whether and to what extent nonhuman animals possess the ability to abstract regularities defining the relation among arbitrary auditory items in a string and to generalize this abstraction to strings of acoustically novel items. In this study we tested these abilities in a songbird (zebra finch) and a parrot species (budgerigar). Subjects were trained in a go/no-go design to discriminate between two sets of sound strings arranged in an XYX or an XXY structure. After this discrimination was acquired, each subject was tested with test strings that were structurally identical to the training strings but consisted of either new combinations of known elements or of novel elements belonging to other element categories. Both species learned to discriminate between the two stimulus sets. However, their responses to the test strings were strikingly different. Zebra finches categorized test stimuli with previously heard elements by the ordinal position that these elements occupied in the training strings, independent of string structure. In contrast, the budgerigars categorized both novel combinations of familiar elements as well as strings consisting of novel element types by their underlying structure. They thus abstracted the relation among items in the XYX and XXY structures, an ability similar to that shown by human infants and indicating a level of abstraction comparable to analogical reasoning.

INTRODUCTION

One of the critical features of language learning is the ability to abstract the grammatical structure from spoken language. Such abstraction allows humans to learn about regularities in their native language and to generalize these regularities to novel input. This ability is examined in a standardized way in artificial grammar learning experiments, in which humans are exposed to strings of meaningless sounds (e.g., arbitrary speech syllables) organized according to a specific grammatical structure. Several studies have shown that the ability to abstract the underlying structure from such stimuli is present in young infants (1–5) in both the acoustic and the visual domain (6–8). This domain generality and its presence at a very early age have given rise to the notion that this cognitive ability may have preceded language evolution and served as a basis for present-day linguistic complexity. If so, it raises the question to what extent this ability is confined to humans or also can be found in nonhuman animals. In this context, comparative studies on nonhuman animals are needed to reveal the level of abstraction they are able to achieve in artificial grammar learning tasks. This information might provide hypotheses about how and why the more complex human grammatical competences have arisen. The current study addresses whether two bird species, the zebra finch and the budgerigar, are able to abstract XYX and XXY grammatical structures consisting of auditory items.

By now, a number of studies on birds (9–16), rats (17–19), and primates (20–22) have used the artificial grammar paradigm to address animal abilities to learn about grammatical structures. Many of these studies used a discrimination task in which the animal has to learn to distinguish two differently structured string sets which are related to differences in reinforcement. However, success in such discrimination learning does not necessarily indicate that the animals learned the underlying structure of the string sets. They might, for instance, have solved the discrimination task by rote memorization of the different strings without detecting any structural relation between them. The critical test, and the hallmark of abstraction as shown by humans, is whether the discrimination between differently structured string sets is maintained in a test in which the novel test strings share no observable physical (acoustical or visual) similarity to the training strings, but only a relational one (23–26), i.e., where the only way to classify the novel stimuli is by the structural similarity between the training and test strings. This ability to maintain the discrimination is nicely illustrated by the classic study by Marcus and colleagues (1), in which infants were familiarized with a series of examples of an XYX structure (with "X" and "Y" items being speech syllables, resulting in strings such as "ga-ti-ga" or "li-na-li"). After familiarization, the

infants were tested with syllables from novel categories (i.e., different consonants and vowels) arranged according to the same structure (e.g., "wo-fi-wo" for the XYX structure) or a different structure (e.g., "wo-wo-fi", the XXY structure). The X and Y items are all drawn from the same pool, and in principle the only cue for discriminating between the novel strings is the similarity or dissimilarity in the arrangement of the items. Whether this ability to discriminate artificial grammar strings by their structure alone is present in nonhuman animals is still an open question, as the experiments that have tested this ability in a variety of mammal species [rhesus monkeys (22), rats (17)] and bird [Bengalese finch (13), zebra finch (9, 15, 16)] species do not provide unambiguous evidence for this ability. The rhesus monkeys (22), for instance, were first habituated to either XXY or XYY strings in which X and Y were various call types. The same call types were used in tests with structure-consistent or structure inconsistent strings, albeit novel exemplars. The monkeys were not tested with novel call types, and hence the generalization of the response could be based on the auditory similarity between novel and familiar string types. Therefore physical (acoustic) similarity, not structural similarity, between training and test items might have guided the responses. The rats continued to discriminate the patterns when the tones were transposed to different frequencies, but it is disputed whether this reflects recognition of the abstract pattern, or recognition of the novel tone sequences as transpositions of familiar ones (27, but also see ref. 28).

The ability of animals to detect grammatical structures has also been tested using an artificial grammar in which several training items belonging to two different categories (A and B) were arranged either as an $(AB)^n$ or an A^nB^n grammar. A series of experiments on tamarins (21), starlings (12), and pigeons and keas (14) demonstrated that all species discriminated among the training sets. However, in order to show what these species learned about the grammatical structure, they were tested with strings consisting of novel items belonging to the same categories as the training items. The tamarins were trained and tested with the A and B items being male and female voices (21); for the starlings, the A and B items were warble and rattle song syllables (12); and for the pigeons and keas, the A and B items were visual tiles showing either a circle or small squares (14). Such test strings could thus always be classified by attending to the acoustic or visual category identity of the training and test items. Only one experiment, in zebra finches, tested whether any species could transfer discrimination between $(AB)^n$ and A^nB^n strings to similarly structured strings consisting of items belonging to novel categories C and D (16). Like the other species, zebra finches could discriminate correctly between test strings with novel items belonging to A and B

categories, but most birds failed to discriminate the two string types when they consisted of C and D element types. There was one notable exception of a bird that did transfer the discrimination to strings with the novel element types. Zebra finches also have been tested using the XYX vs. XXY or XYY grammars discussed above (9, 15, 16). As in previous studies, they had no difficulty in discriminating between the training strings, but when tested with novel combinations of familiar elements or strings with novel element types their responses varied. Some individuals ignored novel stimuli, whilst others generalized based on acoustical similarities between the training and test stimuli (9, 15). Thus, there seems a range in acoustic string discrimination strategies among zebra finches, but there is no clear evidence that they can abstract the underlying structure.

To conclude, the various experiments discussed above either did not address whether nonhuman animals can generalize to novel strings having only a structural similarity to the training strings or have not provided unambiguous evidence that they can do so (24, 26). Therefore, it is still an open question whether they can show the same abstraction shown by infants in the study of Marcus et al. and others (1, 2). Addressing this question is important, not only in the context of grammar learning experiments but also for understanding animal cognition more generally. This type of abstraction is a cognitively advanced ability, comparable to a relational or analogical match-to-sample task as examined in studies on visual pattern abstraction. In a normal match-to-sample task, animals can obtain a reward by choosing from two alternatives the object that matches the one shown previously. However, in a relational match-to-sample task an animal is trained with pairs of stimuli bearing a particular relation to each other, for instance being either the same as (e.g., AA, BB) or different from (e.g., AB) each other. Next they are tested with pairs of novel items that do not share a physical resemblance with the training stimuli, but do share the same underlying structure, such as CC or CD. Such a task thus requires the matching of relations between relations (29–33). Similarly, learning to identify correct sequences of novel auditory items, such as an XYX or XXY structure, requires that the animal first detects that the relation between the X and Y items differs in the XYX and XXY samples and next abstracts this relation to novel samples.

In the current study, we also used the artificial grammar-learning paradigm with XYX and XXY stimuli. One species we used is the zebra finch. Like humans, zebra finches are vocal learners. Their songs consist of ordered, rapidly produced elements. They are a widely used model species for comparative studies on speech and language (34) as well as for auditory learning (26). In addition, we used budgerigars, a representative of another vocal learning clade, the parrots. Budgerigars are vocal learners with more

vocal variation than zebra finches and an open-ended learning capacity. This parrot species produces long and flexible warble songs that vary both in the sequential structure as well as in the shape and duration of elements (35, 36). Even with this strong variation, budgerigars are able to detect minor changes in familiar songs (37). Budgerigars also are frequently used in studies on auditory learning. Although zebra finches and budgerigars have similar hearing abilities, a comparative study showed that budgerigars are better at identifying acoustic stimuli (38). Thus far, budgerigars have not been tested in an artificial grammar learning task, making it of interest to compare them to zebra finches. A comparison of species belonging to these two distinct clades also is of interest because the previously mentioned studies comparing another parrot species, the kea, with pigeons (also a different clade) revealed that these species used very different strategies to distinguish between two string sets consisting of different visual patterns (14, 39). All keas showed the same consistent strategy, but the pigeons showed no consistent pattern at either the species or individual level (39). As mentioned above, previous experiments on zebra finches showed a considerable variety in strategies in distinguishing auditory strings (9, 15, 16) and, as parrots have not yet been tested in an auditory artificial grammar paradigm, our experiment may reveal whether here too the parrot species behaves differently.

Both zebra finches and budgerigars were first trained in a go/ no-go paradigm to discriminate between five XYX and five XXY triplets. These triplets consisted of zebra finch song elements or, in an additional zebra finch experiment, artificial sounds (Fig. S1). After reaching the criterion for reliable discrimination of the training stimuli (Methods), all birds received six different sets of test stimuli, henceforth described as tests 1–6. Test 1 consisted of triplets also used during the training, but now without reinforcement, providing the control condition. The stimuli used in tests 2–5 consisted of items that also occurred in the training stimuli but were arranged in new combinations. Test 6 consisted of triplets involving items that had not been used for the training stimuli. These different tests give rise to different predictions about how the test stimuli should be classified, depending on what the subject learned during training. If the birds use rote memorization of the individual stimuli, all novel stimuli should be considered as more or less equal, irrespective of their structure. On the other hand, if the birds learn to discriminate between the stimuli based on their structure (XYX or XXY), they should classify all test stimuli according to this structure. Finally, if the birds categorize the test triplets according to their physical similarity with the training triplets, there should be variation among the responses over the various tests, because these similarities differ in the different triplet combinations. The results show striking species-specific learning strategies. The zebra finches categorized the test triplets based on their similarity to the training triplets, i.e., they

used the similarity in the ordinal position of the items in the training and the test strings to discriminate the test strings. They did not discriminate among the new triplets. Budgerigars, on the other hand, generalized the discrimination learned in the training strings to new arrangements of familiar sounds and, in an additional experiment, transferred that discrimination to novel items. Therefore they learned the underlying structures of the training strings, thus demonstrating a level of abstraction rarely observed among nonhuman animals.

Results and Discussion: Zebra Finches and Budgerigars

Results: Zebra Finches. When trained with zebra finch song stimuli, the zebra finches learned to discriminate between the two sets of training triplets correctly in, on average, 13,850 trials (SD = 8,124). During the test phase, they correctly discriminated between the XYX and the XXY triplets of test 1 (training triplets; P = 0.014) and test 2 (new combination; P = 0.015; see SI Test Stimuli for stimulus composition and terminology). They did not show discrimination between the XYX and XXY triplets in test 3 (new combination and one new position; P = 0.99), test 4 (new combination and two new positions; P = 0.092), or test 6 (new elements, P = 0.201). Surprisingly, in test 5 (new combination and three new positions) the birds did discriminate between the XYX and XXY structures, but they responded more often to the triplets in the no-go structure than to the triplets in the go structure (P = 0.008). For example, if a bird had been trained to respond when it heard an XXY structure, it would now go more often after hearing an XYX-structured triplet (Fig. 1A). There was no effect of sex of the bird (F = 2.8, P = 0.15), or whether the go structure during training was XYX or XXY (F = 0.33, P = 0.58), or whether the bird had experience with the go/no-go paradigm in a previous experiment (F = 1.95, P = 0.21).

The responses of individual zebra finches varied in the different tests (Fig. 1B). All zebra finches correctly discriminated between the go and no-go stimuli (all test scores fell outside the 95% confidence interval of a binominal distribution, indicated by green dots in Fig. 1B) only in test 1. The other tests showed more variation among the individual birds. In test 3, in particular, one bird responded more to the stimuli with a no-go structure than to the stimuli with the go structure (indicated by a red dot in Fig. 1B), and one zebra finch discriminated the stimuli correctly. None of the other subjects deviated from chance level (black dots in Fig. 1B). When two elements of the triplet were in a position similar to the opposite training structure (test 5), three birds discriminated the stimuli by responding more to the no-go structures. In test 6 none of the birds differed from chance level. Despite this individual variation, Kendall's τb test showed a systematic decrease of the test score over tests 2–5 ($\tau = -0.623$, $P < 0.001$).

A second group of zebra finches trained on triplets consisting of artificial sounds showed similar results (Fig. 1C). They learned to discriminate the training triplets in 13,238 trials (SD = 3,482). Again, they discriminated correctly only in tests 1 and 2 (test 1, $P < 0.01$; test 2, $P < 0.01$). No significant discrimination was seen in tests 3, 4, and 6 (test 3, $P = 0.7$; test 4, $P = 0.17$; test 6, $P = 0.1$). As in the previous group, these zebra finches also discriminated significantly between the XYX and XXY triplets in test 5 by responding more to the no-go items ($P = 0.01$). Again, there was no effect of sex ($P = 0.91$) or of the structure of the go stimuli ($P = 0.78$).

The individual responses of the zebra finches to the artificial stimuli also are very comparable to those of the zebra finches hearing zebra finch song elements. All zebra finches responded more to the go structure than to the no-go structure only in test 1. Again, the greatest variation is shown in test 3, in which two individuals made a correct discrimination and one bird discriminated incorrectly. In test 4, also, one individual made the correct discrimination, and one individual responded more to the no-go structure than to the go structure. These individuals did not discriminate in tests 3 or 5. In test 5, one individual responded more to the test stimulus having the no-go structure. In test 6, the generalization test, one of the zebra finches discriminated between the XYX and XXY triplets incorrectly. As in the previous experiment, these birds also showed a systematic decline in responses to the go- type stimuli over tests 2–5 ($\tau = -0.454$, $P = 0.001$).

Discussion: Zebra Finches. The zebra finches learned to discriminate between the XYX and XXY patterns of both zebra finch song elements and artificial sounds. When the birds heard triplets consisting of novel items, their response rates dropped, and they did not discriminate between the XYX and XXY structures. However, their responses to new combinations of training elements differentiated according to the positional similarities between test and training items. They maintained the correct discrimination between XYX and XXY only in test 2, in which the test strings showed a greater positional similarity to the training strings of the same structure than to the other structure. When the similarities in the positions of elements of the test triplets shifted toward a greater similarity to the triplets of the no-go structure, the zebra finches responded more often to those test triplets. There is a clear overall pattern from responding more to the go structure toward responding more to the no-go structure with increased positional similarity of the test triplets to the no-go training triplets. These results indicate that the zebra finches learned the positions of the elements in the various training stimuli and responded to the test strings based on the similarity in item positions.

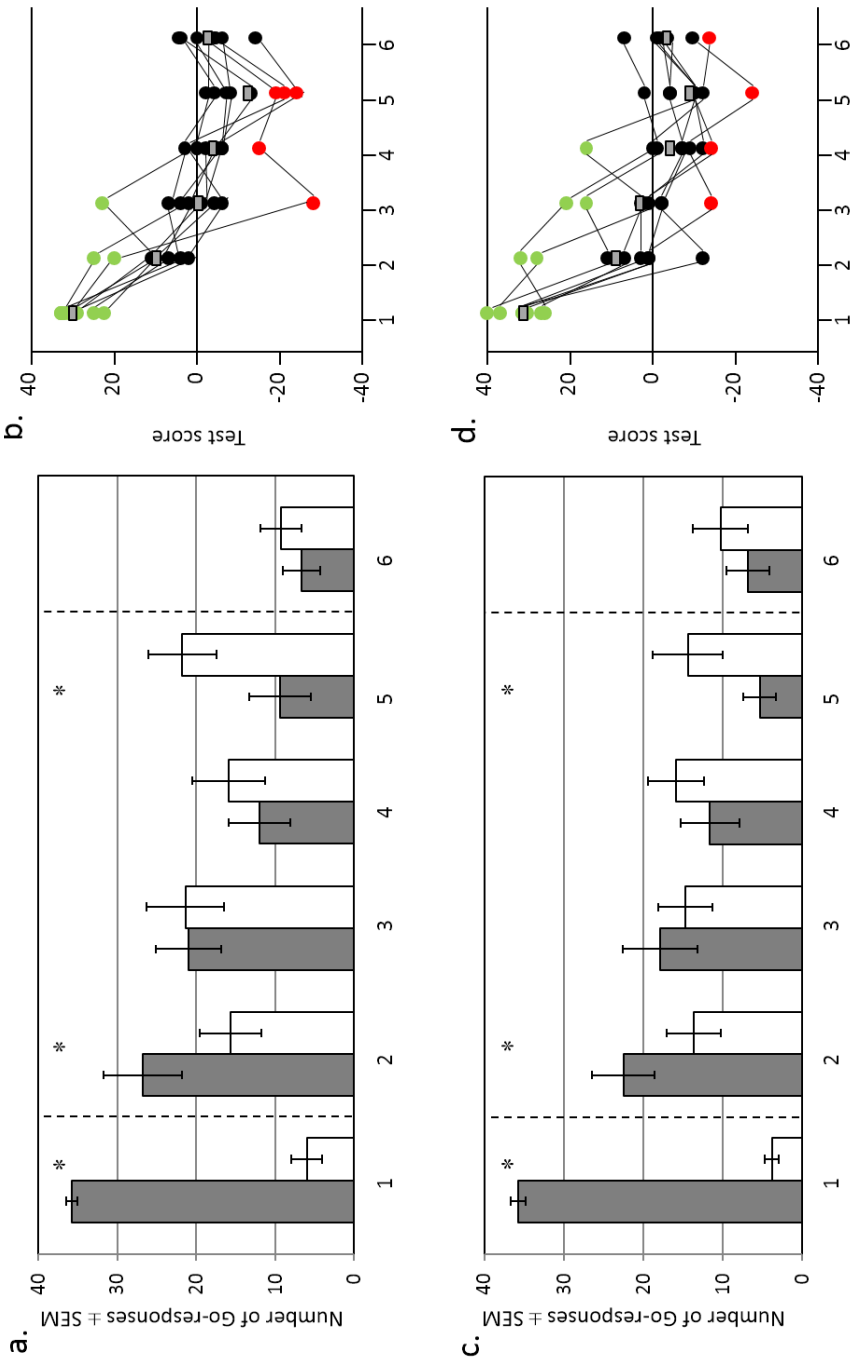


Fig. 1. (A) Average proportional responses (±SEM) in tests 1–6 to the two different structures of the test items for the zebra fiches that heard zebra finch song elements. The grey bars indicate the responses to stimuli with a structure similar to the go stimuli. The white bars indicate the responses to the stimuli with the no-go structure. (B) The test scores of the individual zebra finches that heard zebra finch song elements. A green dot indicates a correct discrimination between the XYX and XYY triplets. Black dots indicate no difference between the responses to XYX and XXY triplets. Red dots indicate a reversed response, in which the structure that was trained as no-go received more responses than the go-structure. Grey blocks show the average responses of the group. (C) Average proportional responses (±SEM) of the zebra finches that heard artificial sounds. (D) The test scores of the individual zebra finches that heard artificial sounds. Group averages are shown by grey blocks. Asterisks in A and C indicate significant differences in response to the XYX and XXY structures.

Results: Budgerigars. The budgerigars were subjected to two consecutive experiments. The first experiment ("mixed test") was a replication of the first zebra finch experiment. The second experiment was a transfer task. After training in the second experiment, the birds were not tested, but the training strings were instantly replaced by novel strings consisting of zebra finch song elements of categories not heard previously. Transfer tests are commonly used to assess whether animals are able to transfer a learned distinction to novel stimuli. With no familiar sounds presented after the transfer, we tested whether the budgerigars were incapable of generalization based on structural similarities or were able to make such generalizations but did not show this ability in the mixed test.

Mixed Test. Like the zebra finches, the budgerigars learned to discriminate between the XYX and XXY triplets during training (mean trials, 7,816 ± 4,845). In contrast to the zebra finches, the budgerigars consistently discriminated between the test strings with familiar items according to their structural similarity to the training strings (test 1, $P = 0.031$; test 2, $P = 0.035$; test 3, $P = 0.041$; test 4, $P = 0.035$; test 5, $P = 0.043$). In test 6, generalization to novel elements, the budgerigars responded hardly at all and did not discriminate between test strings ($P = 0.31$) (Fig. 2A). Here, also, there was no effect of sex of the bird ($t = -6.01$, $P = 0.31$) or of the go structure during training ($t = -5.37$, $P = 0.24$). The majority of the individual budgerigars followed the pattern seen at the group level (Fig. 2B). All birds discriminated between the stimuli in test 1, and at least one-half of the birds made a significantly correct discrimination in tests 2–5. In test 6 none of the birds discriminated between the test triplets. One bird showed a consistently high correct discrimination in tests 1–5, but its performance also dropped to the chance level in test 6. Nevertheless, the figure shows that if a bird did respond in test 6, that response was more often to a string with the grammatically correct structure. There was no significant relation between the test score and the test number over tests 2–5 ($\tau = -0.096$, $P = 0.55$), indicating that the budgerigars' responses are not affected by the similarity in item positions in the training and test strings.

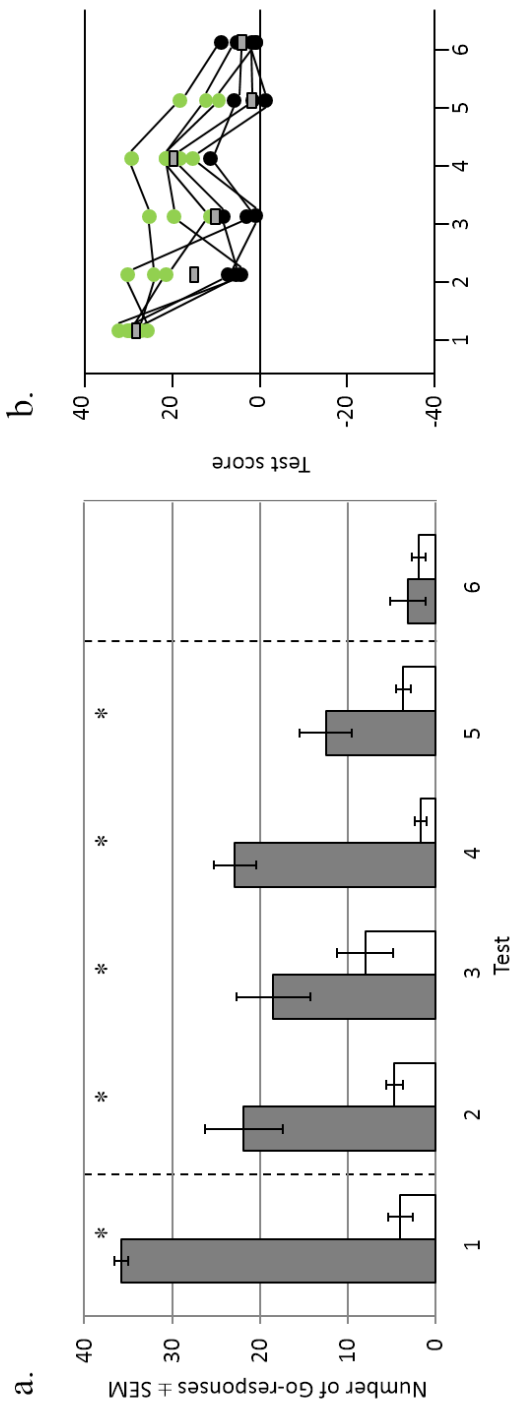


Fig. 2. (A) The average proportion of responses of the budgerigars in the different tests. The grey bars indicate the responses to stimuli with a structure similar to the go stimuli. The white bars indicate the responses to stimuli with the "no-go" structure. Asterisks indicate significant differences in response to the XYX and XXY structures. (B) The test scores of the individual budgerigars. A green dot indicates a correct discrimination between the XYX and XYY triplets. Black dots indicate no difference between the responses to XYX and XXY triplets. Grey blocks show the average test score of the group with the same color coding.

Transfer task. All birds learned to discriminate between the go and no-go items during the first training phase and reached the discrimination criterion faster than in the previous experiment (mean ± SD mixed test = 7,816 ± 4,845; mean ± SD transfer task = 6,483 ± 4,491; $P < 0.01$). Of the six budgerigars, five retained their discrimination after the transfer to new song elements (Fig. 3) as shown by their discrimination ratios (correct go responses/ all go responses) for the transfer sets. Transfer set 1 consisted of the first responses to each of the 12 XYX and 12 XXY stimuli, transfer set 2 consisted of all responses to the second time the subjects heard the stimuli, and so forth. The five budgerigars discriminated the two stimulus structures significantly above chance level for the first transfer set and remained above chance for all subsequent sets (individual budgerigar P2: all $P < 0.04$; budgerigar P4: all $P < 0.001$; budgerigar P8: all $P < 0.02$; budgerigar P9: all $P < 0.01$; and budgerigar P11: all $P < 0.01$). One individual did not perform better than chance after the transfer, although he also learned to discriminate between the training items by the fifth time he encountered the stimuli (budgerigar P7: test set 1–4, all $P > 0.19$; set 5, $P < 0.01$) (Fig. 3).

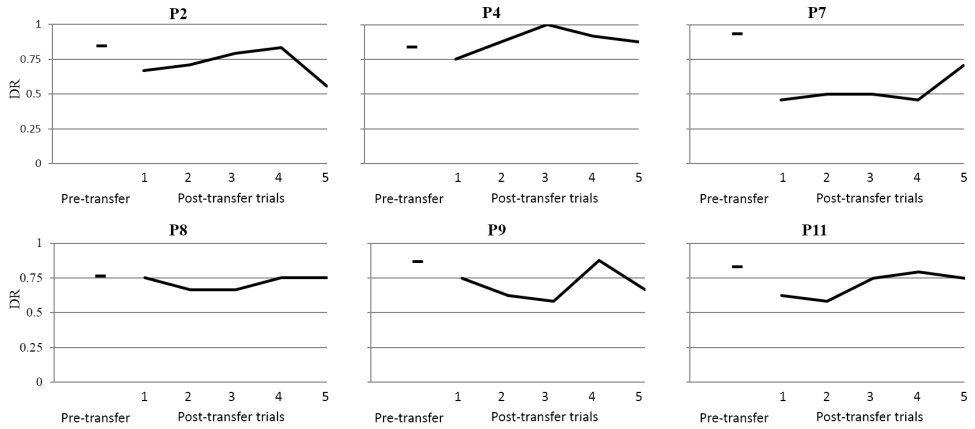


Fig. 3. The discrimination ratio per set of test trials (each set consists of 12 different go and 12 different no-go triplets) per individual. The short line indicates the average discrimination ratio (DR) of the last 100 trials of the training. The line shows the responses to the first five trial sets (1–5) after being transferred to new stimuli. Five of six budgerigars discriminated above chance level immediately after the transfer.

Discussion: Budgerigars. The budgerigars successfully categorized triplets in the mixed test by their structure as long as the items in these triplets were familiar. They hardly responded to triplets consisting of new items and showed no significant difference in response to those triplets. When they did respond, however, they responded more toward triplets that were similar in structure to the go stimuli. Their

correct categorization of triplets in which all items were in new positions shows that they generalized based on the structure of the training triplets. If they had not generalized the structure, their responses would have been at chance level; if they had generalized based on positional similarities between training and test items, their responses would have been more similar to those of the zebra finches.

The low number of responses to the stimuli with new items in test 6 might reflect a weakness of the test procedure of the go/no-go task that has been noted before (9, 15, 16, 40). Because only the go training stimuli (consisting of familiar items) provided a reward, the budgerigars might quickly have learned to ignore any stimulus containing novel items. However, although they ignored the strings consisting of novel items in the mixed test, they nevertheless noted the structure of these strings, because, averaged over the first encounters with the novel training strings of the transfer task, five of the six budgerigars immediately discriminated correctly between XYX and XXY triplets. These responses demonstrate an immediate generalization of the structural information provided by the training stimuli.

GENERAL DISCUSSION

Our results show that in the same learning task zebra finches used positional information to discriminate between string sets, whereas budgerigars used the grammatical structure underlying the training strings to categorize new strings. The abstraction shown by the budgerigars requires learning the relations between items in a string. The budgerigars first had to learn that the relation between the items differed in the XYX- and the XXY-structured triplets and then had to link these differences to their actions and to the received feedback. When hearing unfamiliar sounds, they needed to perceive the relations between these new items and compare these relations to the structural relations between the training items to give the appropriate response. Nonhuman species' ability for such abstraction has been debated and, until now, had not been clearly shown (see Introduction and ref. 24). Several studies have shown that animals can discriminate between artificially constructed sound strings (e.g., 8–10, 12, 14, 15). However, this discrimination usually is based on attending to specific, familiar sound features, like repetitions of items belonging to a familiar category (9, 15, 16).

The results of the budgerigars were obtained by using a different experimental paradigm (go/no-go) from the one most commonly used in infant studies (familiarization). Nevertheless, the abstraction shown by the budgerigars seems comparable to that observed by Marcus et al. (1) in human infants at the age of 7

months old (1), who were able to learn structures such as XYX vs. XXY. However, as noted for rule learning in humans (e.g., refs. 41–43), sets of strings can be discriminated in many ways, ranging from rote memorization of the different strings to using the abstract structure underlying a set. Similarly, the training stimuli used in our present experiment allowed discrimination based on structure (XYX vs. XXY, or every other diagnostic form, such as a repetition of X) as well as discrimination based on memorization of element positions. Interestingly, infants do not always discriminate XYX and XXY strings based on their structure but also may use positional information that is present in training strings. Gerken (41) presented two groups of infants with different subsets of triplets from the stimuli of Marcus et al. (1). One group heard four XXY triplets in which both the X and the Y syllable were different in each triplet. The second group heard a subset of four triplets in which the Y syllable was constant (always "di"). Infants in the first group generalized the structure of the triplets to novel exemplars, but infants in the second group generalized only to novel exemplars with the di syllable in the final position. Thus, depending on the type of information presented during familiarization, infants attended either to the grammatical structure or to the ordinal position of specific items in the triplets. In our experiment, in which the training stimuli could be discriminated either on structure or on item positions, the zebra finches are clearly biased to memorize the ordinal positions of familiar elements, while the budgerigars attended to the underlying structures. The behavior of the two species thus reflects two different strategies, both of which are present in human infants.

The positional learning observed in the zebra finches is a strategy commonly seen when animals are trained to discriminate between strings of visual or acoustic stimuli. Zebra finches are known to use the positional information in strings of song elements, although they also can use transitional information (44). Other animals also use positional information in auditory and nonauditory string-learning tasks. For instance, when pigeons are trained to respond to sequences with an A-B-C-D structure (the letters depicting different image categories), they respond more quickly when they see new strings in which the category items are in the original ordinal position than when they see strings in which the category items are switched (e.g., A-C-B-D) (45, 46). A similar mapping of items to positions in both string-discrimination and food-localization paradigms has been noted in several primate species (20, 47, 48). While the positional learning as observed in the zebra finches is thus not uncommon, the detailed memory for item positions that they demonstrate in the current experiment is impressive, as they kept track of item positions over a set of ten training strings.

Learning about abstract structures and relations, as shown by the budgerigars, is rare in the animal kingdom. The cognitive process underlying abstraction is considered a higher-order cognitive ability as it requires an individual to ignore the specific features of individual items and to attend to the global, structural features, in this case of the sound strings. This ability is analogous to the cognitive skills used in relational match-to-sample tasks. As described in the Introduction, success in these tasks is possible only when the subject attends to the relation between items (e.g., whether two items are the same or different). The subject must transfer the learned relations between familiar items to novel ones during testing (30–33), comparable to the requirements for success in our artificial grammar learning task. The ability to perform such a relational match-to-sample task long was thought to be unique for humans; it also has been demonstrated in some great ape species (30) and recently has been demonstrated in crows and amazon parrots (32, 33, but also see ref. 49), but so far only in the visual domain. Our findings suggest that the species that succeeded in a relational match-to-sample task, might also succeed in artificial grammar tasks such as the current one.

While our experiment reveals a clear species difference in the strategies used to discriminate the string sets, the cause of this difference is not clear. It is unlikely that the observed difference is due to the nature of the sound items. Zebra finches trained and tested with unfamiliar, nonsong stimuli also used the positional learning strategy, so the use of this strategy is not linked to the items being species-specific sounds. For the budgerigars, the zebra finch song elements were equally as unfamiliar as the artificial sounds were to the zebra finches. It is more likely that the difference in strategy is based on a species- or clade-specific difference. The two species have different singing styles: The budgerigars are open-ended vocal learners, whereas vocal learning in zebra finches is limited to a sensitive phase during development. Whether or how such differences might be responsible for the observed difference in discrimination is unclear. A likely alternative explanation may be that the budgerigars are more cognitively advanced in solving complex learning tasks or in abstraction in general. This explanation also may be supported by the previously mentioned study by Dent et al. (38), which showed that budgerigars are better than zebra finches at identifying complex acoustic stimuli (although that study did not address the budgerigars' ability for abstraction). In our experiment, also, the budgerigars required fewer trials to reach the training criterion. Our findings thus confirm that different bird species trained and tested on the same stimulus sets can behave quite differently, as was observed for the keas and pigeons in a visual grammar learning task (39). In that study pigeons and keas both attended to local features of the training stimuli; the species differed in which features were used and the consistency with which a specific strategy was used among individuals. In the current study, the difference between

species seems to be a more fundamental one, with both zebra finches and budgerigars being consistent within the species, but one attending to local features and the other to a higher-order regularity. It is obvious that further comparative work is needed to reveal the origin and scope of the cognitive differences between the various bird species. As noted above, both the parrot and the songbird clades contain cognitively advanced species (50, 51) and hence provide ample opportunity for such studies.

CONCLUSIONS

We have demonstrated the ability of a vocal learning avian species, the budgerigar, to learn about the abstract relations between items in an acoustic string. Our results indicate that the cognitive system necessary to detect grammatical regularities in sound strings is not unique to humans or to having language. Instead, it seems to be an ability that is uncommon throughout the animal kingdom, and may indicate a case of parallel evolution between humans and at least one nonhuman animal species.

METHODS

This research was approved by the Leiden Committee for Animal Experimentation (application numbers 12214 and 12215).

ZEBRA FINCHES

Subjects

Sixteen zebra finches [eight males, eight females, 202–482 d post hatching (dph)] from the breeding colony of Leiden University were used, equally divided between the two experiments. Six birds had been trained previously in a go/no-go paradigm but not in a comparable learning task, and the previous training had been with human speech sounds rather than song elements. Nevertheless, we factored this training as a potential effect in the analysis. Before the experiment, the zebra finches lived in single-sex groups on a 13.5-h/10.5-h light/dark schedule at 20–22 °C. Food, water, grit, and cuttlebone were available ad libitum.

Apparatus

All experiments were conducted in an operant conditioning cage (zebra finches: 70 cm long × 30 cm wide × 45 cm high; budgerigars: 70 cm long × 60 cm wide × 60 cm

high). Each operant cage was in a separate sound-attenuated chamber and was illuminated by a fluorescent tube that emitted a daylight spectrum on a 13.5-h/10.5-h light/dark schedule. A speaker (Vifa 10BGS119/8) was located 1 m above the center of the cage. The cage walls were made from wire mesh except for the plywood back wall which supported two pecking keys with LED lights. A food hatch, easily accessible to the birds, was located between these two keys. Pecking the left key (sensor 1) elicited a stimulus and illuminated the LED light of the key on the right (sensor 2). Depending on the sound, the bird had to peck sensor 2 (the go response) or withhold its response (the no-go response). A correct pecking response resulted in access to food for 10 s, and an incorrect peck led to 15 s of darkness.

Stimuli

The stimuli consisted of three sound items concatenated in triplets according to either an XYX or an XXY structure. One group of zebra finches received elements from natural zebra finch songs, which were ramped (3 ms) and rms (0.1) equalized in PRAAT (version 5356, freeware available at www.praat.org). The other group received artificially created sounds, such as a clap, a trumpet, white noise, or pure tones, also ramped and rms equalized. The frequency of these sounds was modified so that the peak frequencies fell within the hearing range of the zebra finches. Durations were adjusted to be in the same range as the zebra finch song elements. Triplets were created from either three zebra finch song elements or three artificial sounds, with 40 ms of silence between the sound items and 50 ms silence at the beginning and the end of the triplet. For each bird, 10 triplets were created for the training phase; five of these triplets followed the XYX structure, and five triplets followed the XXY structure (Fig. S1 and Table S1). The test triplets were created in a fashion similar to the training triplets but consisted of new combinations of the elements and also novel elements not heard before (Table S2). We created six different types of test stimuli, which were presented to the birds in random order. Details of the composition of the test stimuli are provided in SI Test Stimuli.

Procedure

Before the experimental training phase, each bird was conditioned to become familiar with the go/no-go paradigm, using a conspecific song as the go stimulus and a pure tone as the no-go stimulus. When the birds reached our standard discrimination criterion of more than 75% correct responses to the go stimuli and fewer than 25% incorrect responses to the no-go stimuli on two consecutive days, they proceeded to the training phase. In the training phase all birds received five different go triplets and five no-go triplets (Table S1). For one-half of the birds, the triplets with an XYX

structure were the go triplets; for the other birds the triplets with an XXY structure were the go triplets. When the birds reached our standard discrimination criterion for the five stimuli on average for three consecutive days, they were subjected to an accustoming phase before the test phase started. In this accustoming phase all stimuli occurred both with and without feedback at an 80–20% ratio. When the birds' performance returned to the criterion level (usually within 1 or 2 d), they proceeded to the test phase. During the test phase a peck on sensor 1 was followed by a training triplet in 80% of the cases, and the subsequent behavior of the bird was reinforced. In 20% of the cases the bird heard a test triplet after pecking sensor 1. In these cases the response of the subject was not reinforced. The test phase lasted until each test triplet appeared 40 times. All test triplets were mixed at random with the training triplets.

Analyses

The analyses were conducted on an individual level as well as on a group level. First, we calculated a test score for each bird for each test. This test score was calculated as the number of correct responses to the test triplets structured like the go training triplets minus the incorrect responses to the test triplets structured like the no-go training triplets consisting of the same elements (e.g., DCD versus DDC): test score = response to go triplet − response to no-go triplet. Because all birds were tested with 40 go and 40 no- go triplets for every test condition, their test scores could range from +40 to −40, with +40 being perfect discrimination and 0 being performance at chance level. The individual analyses were done with a binominal test with a Holm correction, in which this test score was measured against a binominal distribution with a success probability of 0.5. When the individual test score was within the 95% confidence interval of the binominal distribution, we report this result as having $P > 0.05$. The group analyses were done with a linear mixed model, based on the test scores. We tested for an effect of sex and go- item structure (XYX or XXY). To analyze the difference in response per test on a group level, we performed a paired Wilcoxon signed rank test between the number of responses to the test stimuli with the go structure and the number of responses to the test stimuli with the no-go structure. We examined whether there was a systematic change in the test scores over tests 2–5 using Kendall's τb test.

BUDGERIGARS

Subjects

Six budgerigars (two males, four females, 263–389 dph) were used in both the mixed test and the transfer experiments. They came from different breeders in the

Netherlands and were individually housed on a 13.5-h/10.5-h light-dark schedule at 20–22 °C before the experiment. Food, water, grit, and cuttlebone were available ad libitum.

Stimuli

The mixed test was conducted with the same zebra finch song stimuli in the same configurations used for the zebra finches. However, we extended the pauses between consecutive elements to 100 ms, because a pilot experiment suggested that the budgerigars found it harder to discriminate the rapid succession of the zebra finch song elements.

Training stimuli for the transfer experiment were constructed from the zebra finch song elements heard by the budgerigars in the first experiment. With these eight elements we created a balanced set of eight XYX and eight XXY triplets (Table S3). Four new types of zebra finch song elements were used to create 12 XYX and 12 XXY stimuli for the second training phase (Table S4). These new elements had not been heard previously and were the same for all budgerigars.

Procedure

The mixed test was conducted in an identical fashion to the tests with the zebra finches. The procedure of the transfer test was as follows. Before the start of the training (between 203 and 404 d after the end of the previous experiment), the budgerigars were again conditioned in the go/no- go set-up. A warbled song from a budgerigar unknown to the subject birds served as the go stimulus, and a pure tone served as the no-go stimulus. When the birds reached the standard learning criterion on three consecutive days, they proceeded to the first training phase in which the birds were trained to discriminate between XYX and XXY triplets. Each individual bird received the same go-structure and reinforcement as during the previous experiment. This phase lasted until each individual bird reached the standard criterion for three consecutive days. The subject was next immediately transferred to a second training phase. In this second phase, the budgerigars were presented with 12 go and 12 no-go triplets, each of which consisted of novel elements. Reinforcement in this second training phase was identical to the first training phase of the experiment: food was available after a correct response and the lights switched off after an incorrect response. After 2,000 trials in the second training phase, the birds were returned to their home cages, regardless of their performance in this phase.

Analysis

The results from the mixed test were analyzed in a similar fashion as the zebra finch results. The results from the transfer test were calculated as discrimination ratios (correct go responses/all go responses). These ratios were analyzed for deviations from chance level with a generalized linear mixed model. Because the birds received reinforcement during both the first and the second training phase of the transfer test, only the data from the first five sets of trials after the transfer were used for the analysis.

Acknowledgements

We thank Andreea Geambasu, Claartje Levelt, Raquel G. Alhama, and Jelle Zuidema for suggestions on the experiment and help with the stimuli; and Guus van der Ven, Arne Dits, Ruth van Koppen, Cas Retel, Koen Stam, and Sophie Verstraeten for help with the practical work. This research was supported by Netherlands Organization for Scientific Research- Humanities Grant 360.70.452.

SUPPORTING INFORMATION

SI Test Stimuli

Test 1 consisted of the triplets also used during the training but used in test 1 without reinforcement, providing the control condition. Test stimuli used in tests 2–5 consisted of elements that also occurred in the training stimuli but arranged in new combinations. Test 6 consisted of triplets involving element types that had not been used for the training stimuli. The similarity to the training stimuli differs among the different combinations of test triplets. For example, test 2 consisted of the triplets BDB and BBD. Both triplets thus consist of two B elements and one D element. Among the training triplets, there are two stimuli also containing two B elements: BFB and BBF. In addition, there are six triplets containing one B element or one D element (ADA, AAD, DFD, DDF, EBE, and EEB). The BDB and BBD triplets are composed of the same elements, which appear equally often in the go and no-go training triplets. Thus, if the preference expressed in the test was based on the number of elements shared by the test and the training triplets, the birds should show equal responses to BDB and BBD. However, this prediction changes if the birds attend to the position of multiple shared elements. BDB and BFB share two B elements, and these elements are in identical positions. Similarly, BBF and BBD share two identically positioned B elements. If test triplets were judged based on the positional similarities in stimuli that share multiple elements with the training items, BDB should be treated as being more similar to an XYX triplet, and BBD should be treated as being as more similar to a XXY triplet. Like BDB, both ADA and DDF have D in the middle position. BBD, AAD, and DFD all have D in the final position. Thus, with respect to the D element, BDB is similar to an XYX training stimulus, and BBD is similar to an XXY training stimulus. Hence, overall, the positional similarities of the elements in the BDB and BBD stimuli are closer to training triplets that have the same structure as these test triplets than to training triplets with the opposite structure (Table S2). This balance shifts gradually over tests 3–5. The FEF and FFE stimuli of test 5 never share more than one element with the training stimuli, and no training stimuli use two F elements. With respect to the positional similarity, FEF shares the F element in the final position with DDF and BBF and shares the E element in the middle position with EEB. The FFE triplet shares an F in the middle position with BFB and DFD and an E in the final position with EBE. Hence, each of these test triplets shares the position of two elements with stimuli belonging to the other structural category than the test stimulus itself, but they share no positional similarity with training stimuli from their own structural category. For tests 3 and 4 the situation is intermediate. Table S2 presents the overview of the positional similarities between test and training triplets.

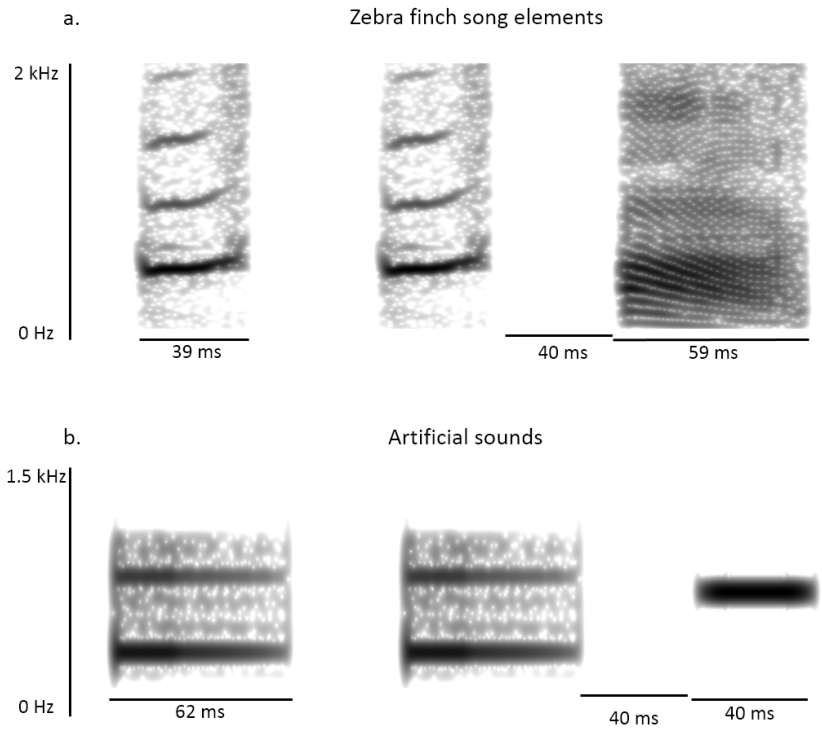


Fig. S1. Examples of XXY triplets concatenated from zebra finch song elements (A) and artificial sounds (B). The lengths of the elements and pauses are indicated below the figure.

Table S1. Representation of the training triplets. The letters depict different zebra finch song elements. These elements differed for each individual, i.e., the A and D elements used to make the ADA triplet for one bird were from different element categories than the A and D elements used to make the ADA triplet for another bird.

XYX	XXY
ADA	AAD
BFB	BBF
CAC	CCA
DFD	DDF
EBE	EEB

Table S2. Representation of the test triplets with information on the number of elements in each test triplet that are in the same position as the training triplets of both grammars.

Test	Stimuli		Positions shared with *same* structure triplets			Positions shared with *other* structure triplets		
	XYX	XXY	n	With training triplet		n	With training triplet	
1- Training Triplets	CAC	CCA	3	CAC	CCA	2	CCA AAD	CAC ADA
	EBE	EEB	3	EBE	EEB	2	EEB BBF	EBE BFB
2- New Combination	BDB	BBD	3	BFB ADA	BBF AAD	3	BBF DDF EEB	BFB EBE DFD
3- New Combination & 1 New Position	DCD	DDC	2	DFD	DDF	3	DDF CCA AAD	DFD ADA CAC
4- New Combination & 2 New Positions	FAF	FFA	1	CAC	CCA	2	AAD BBF	BFB ADA
5- New Combination & 3 New Positions	FEF	FFE	0			2	EEB DDF	DFD EBE
6- New Elements	GHG	GGH	0			0		
	HGH	HHG	0			0		

The column "Stimuli" shows which triplets were used in the different tests. n1 shows the number of elements that the test strings shared with the training strings of the same grammar structure. n2 shows the number of elements shared with the training strings of the alternative grammar structure. The table also shows the training triplets in which these elements occurred (relevant elements are underlined). In test 3, 4, and 5 there are more elements sharing positions with the no-go structure than with the go structure; this can explain why these stimuli are preferred over those that are structurally similar to the training stimuli.

Table S3. Abstract representation of the 16 training triplets used for the transfer experiment with the budgerigars. The letters depict song elements from different zebra finches; these elements were different for each individual.

Training	
XYX	**XXY**
ADA	AAD
BFB	BBF
CAC	CCA
DCD	DDC
EBE	EEB
FHF	FFH
GEG	GGE
HGH	HHG

Table S4. Abstract representation of the triplets in the second training phase of the transfer task. The letters depict different zebra finch song elements, which were novel to the subjects.

Test	
XYX	**XXY**
IJI	IIJ
IKI	IIK
ILI	IIL
JIJ	JJI
JKJ	JJK
JLJ	JJL
KIK	KKI
KJK	KKJ
KLK	KKL
LIL	LLI
LJL	LLJ
LKL	LLK

REFERENCES

1. Marcus GF, Vijayan S, Bandi Rao S, Vishton PM (1999) Rule learning by seven-month- old infants. Science 283(5398):77–80.

2. Gomez RL, Gerken L (1999) Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70:109–135.

3. Marcus GF, Fernandes KJ, Johnson SP (2007) Infant rule learning facilitated by speech. Psychol Sci 18(5):387–391.

4. Pons F, Toro JM (2010) Structural generalizations over consonants and vowels in 11- month-old infants. Cognition 116(3):361–367.

5. Gerken L, Knight S (2015) Infants generalize from just (the right) four words. Cognition 143:187–192.

6. Saffran JR, Pollak SD, Seibel RL, Shkolnik A (2007) Dog is a dog is a dog: Infant rule learning is not specific to language. Cognition 105(3):669–680.

7. Frank MC, Slemmer JA, Marcus GF, Johnson SP (2009) Information from multiple modalities helps 5-month-olds learn abstract rules. Dev Sci 12(4): 504–509.

8. Gebhart AL, Newport EL, Aslin RN (2009) Statistical learning of adjacent and non-adjacent dependencies among nonlinguistic sounds. Psychon Bull Rev 16(3):486–490.

9. Chen J, van Rossum D, ten Cate C (2015) Artificial grammar learning in zebra finches and human adults: XYX versus XXY. Anim Cogn 18(1):151–164.

10. Comins JA, Gentner TQ (2014) Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. Anim Cogn 17(5):1023–1030.

11. Comins JA, Gentner TQ (2013) Perceptual categories enable pattern generalization in songbirds. Cognition 128(2):113–118.

12. Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. Nature 440(7088):1204–1207.

13. Seki Y, Suzuki K, Osawa AM, Okanoya K (2013) Songbirds and humans apply different strategies in a sound sequence discrimination task. Front Psychol 4:447.

14. Stobbe N, Westphal-Fitch G, Aust U, Fitch WT (2012) Visual artificial grammar learning: Comparative research on humans, kea (Nestor notabilis) and pigeons (Columba livia). Philos Trans R Soc Lond B Biol Sci 367(1598): 1995–2006.

15. van Heijningen CAA, Chen J, van Laatum I, van der Hulst B, ten Cate C (2013) Rule learning by zebra finches in an artificial grammar learning task: Which rule? Anim Cogn 16(2):165–175.

16. van Heijningen CAA, de Visser J, Zuidema W, ten Cate C (2009) Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. Proc Natl Acad Sci USA 106(48):20538–20543.

17. Murphy RA, Mondragón E, Murphy VA (2008) Rule learning by rats. Science 319(5871): 1849–1851.

18. de la Mora DM, Toro JM (2013) Rule learning over consonants and vowels in a non- human animal. Cognition 126(2):307–312.

19. Toro JM, Trobalón JB (2005) Statistical computations over a speech stream in a rodent. Percept Psychophys 67(5):867–875.

20. Endress AD, Carden S, Versace E, Hauser MD (2010) The apes' edge: Positional learning in chimpanzees and humans. Anim Cogn 13(3):483–495.

21. Fitch WT, Hauser MD (2004) Computational constraints on syntactic processing in a nonhuman primate. Science 303(5656):377–380.

22. Hauser MD, Glynn D (2009) Can free-ranging rhesus monkeys (Macaca mulatta) extract artificially created rules comprised of natural vocalizations? J Comp Psychol 123(2):161–167.

23. ten Cate C, van Heijningen C, Zuidema W (2010) Reply to Gentner er al.: As simple as possible, but not simpler. Proc Natl Acad Sci USA 107(16):E66–E67.

24. ten Cate C, Okanoya K (2012) Revisiting the syntactic abilities of non-human animals: Natural vocalizations and artificial grammar learning. Philos Trans R Soc Lond B Biol Sci 367(1598):1984–1994.

25. Beckers GJL, Bolhuis JJ, Okanoya K, Berwick RC (2012) Birdsong neurolinguistics: Songbird context-free grammar claim is premature. Neuroreport 23(3):139–145.

26. ten Cate C (2014) On the phonetic and syntactic processing abilities of birds: From songs to speech and artificial grammars. Curr Opin Neurobiol 28:157–164.

27. Corballis MC (2009) Do rats learn rules? Anim Behav 78(4):E1–E2.

28. Mondragon E, Murphy RA, Murphy VA (2009) Rats do learn XYX rules. Anim Behav 78(4):E3–E4.

29. Herrnstein RJ (1990) Levels of stimulus control: A functional approach. Cognition 37(1-2):133–166.

30. Thompson RKR, Oden DL (2000) Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. Cogn Sci 24(3): 363–396.

31. Vonk J, Povinelli DJ (2006) Similarity and difference in the conceptual systems of primates: The unobservability hypothesis. Comparative Cognition, eds Zentall T, Wasserman EA (Oxford Univ Press, Oxford, UK), pp 363–387.

32. Smirnova A, Zorina Z, Obozova T, Wasserman E (2015) Crows spontaneously exhibit analogical reasoning. Curr Biol 25(2):256–260.

33. Obozova T, Smirnova A, Zorina Z, Wasserman E (2015) Analogical reasoning in amazons. Anim Cogn 18(6):1363–1371.

34. Doupe AJ, Kuhl PK (1999) Birdsong and human speech: Common themes and mechanisms. Annu Rev Neurosci 22:567–631.

35. Tu HW, Smith EW, Dooling RJ (2011) Acoustic and perceptual categories of vocal elements in the warble song of budgerigars (Melopsittacus undulatus). J Comp Psychol 125(4):420–430.

36. Farabaugh SM, Brown ED, Dooling RJ (1992) Analysis of the warble song of the budgerigar (Melopsittacus undulates). Bioacoustics 4(2):111–130.

37. Tu HW, Dooling RJ (2012) Perception of warble song in budgerigars (Melopsittacus undulatus): Evidence for special processing. Anim Cogn 15(6): 1151–1159.

38. Dent ML, Welch TE, McClaine EM, Shinn-Cunningham BG (2008) Species differences in the identification of acoustic stimuli by birds. Behav Processes 77(2):184–190.

39. Ravignani A, Westphal-Fitch G, Aust U, Schlumpp MM, Fitch WT (2015) More than one way to see it: Individual heuristics in avian visual computation. Cognition 143:13–24.

40. Spierings MJ, ten Cate C (2014) Zebra finches are sensitive to prosodic features of human speech. Proc Biol Sci 281(1787):20140480.

41. Gerken L (2006) Decisions, decisions: Infant language learning when multiple generalizations are possible. Cognition 98(3):B67–B74.

42. Aslin RN, Newport EL (2014) Distributional language learning: Mechanisms and models of category formation. Lang Learn 64(Suppl 2):86–105.

43. Kovacs AM (2014) Extracting regularities from noise: Do infants encode patterns based on same and different relations? Lang Learn 64(2):65–85.

44. Chen J, ten Cate C (2015) Zebra finches can use positional and transitional cues to distinguish vocal element strings. Behav Processes 117:29–34.

45. Scarf D, Colombo M (2011) Knowledge of the ordinal position of list items in pigeons. J Exp Psychol Anim Behav Process 37(4):483–487.

46. Scarf D, Colombo M (2010) Representation of serial order in pigeons (Columba livia). J Exp Psychol Anim Behav Process 36(4):423–429.

47. Drucker CB, Brannon EM (2014) Rhesus monkeys (Macaca mulatta) map number ontospace. Cognition 132(1):57–67.

48. Terrace HS (2005) The simultaneous chain: A new approach to serial learning. Trends Cogn Sci 9(4):202–210.

49. Vonk J (2015) Corvid cognition: Something to crow about? Curr Biol 25(2):R69–R71. 50.

50. Péron F, Rat-Fischer L, Lalot M, Nagle L, Bovet D (2011) Cooperative problem solving in African grey parrots (Psittacus erithacus). Anim Cogn 14(4):545–553.

51. Péron F, Rat-Fischer L, Lalot M, Nagle L, Bovet D (2011) Cooperative problem solving in African grey parrots (Psittacus erithacus). Anim Cogn 14(4):545–553.

52. Taylor AH, Elliffe D, Hunt GR, Gray RD (2010) Complex cognition and behavioural innovation in New Caledonian crows. Proc Biol Sci 277(1694): 2637–2643.



8

THESIS SUMMARY & GENERAL DISCUSSION



THESIS SUMMARY & GENERAL DISCUSSION

Language is unique to humans, but the underlying mechanisms to produce and perceive language are not all uniquely human. Shared mechanisms did not evolve specifically for language and might be at the roots of the evolution of language as we have it today. In this thesis, I presented several studies on the cognitive mechanisms related to grammar, rhythm and prosody perception in two bird species. Here I will summarize these results in two sections: *Perception of phonological patterns*, which includes chapter 2, 3 and 4, and *Perception of syntactic patterns*, which includes chapter 5 and 7.

Perception of phonological patterns

Language is produced as strings of syllables that together form words and sentences. These sound strings are not segmented into words as clearly as written language is; pauses in speech can occur both between and within words. Infants can still use these pauses when segmenting sentences into words, but also use other cues, such as the transitional probabilities between syllables (1-3). Transitional probabilities are the likelihood that two syllables belong together, based on how often they occurred together. Syllables that are adjacent in a word always occur together when that word is produced, syllables of two different words occur together less often. Furthermore, words have prosodic patterns that are typical for a language. Word onsets are often louder and higher than the offset, whilst in many languages the offset of a word is elongated. Humans have a perceptual bias to group a string of sounds alternating in pitch, amplitude or duration into short units with similar prosodic patterns as words (4, 5).

String segmentation. In CHAPTER 2 we studied how zebra finches segment strings of song elements into shorter chunks. When zebra finches are learning to sing, they copy chunks of their fathers' song and may also insert chunks of songs from other males or improvised elements (6, 7). The chunks that are copied are often separated by slightly longer pauses in the original song (7). This indicates that zebra finches may have a natural tendency to segment strings. The question remains how they do this and in which way this is comparable to the string segmentation strategies of human infants. In this study the zebra finches were trained to discriminate between two long element strings. For one group, the strings could be segmented based on the co-occurrence of the elements. In the second group, pauses were inserted between the co-occurring element groups. The results show that, unlike human infants, zebra finches do not use the co-occurrence of elements in a string to learn the segments.

They can, however, use elongated pauses between the chunks to memorize which elements belong together, similarly to human infants (8).

Perceptual grouping. CHAPTER 3 tested whether zebra finches possess similar perceptual biases to humans to group strings of tones alternating in pitch or duration. Perceptual grouping means that a long string of equally spaced items is not perceived as a continuous sequence of items, but rather as a concatenation of shorter groups of sounds. In the acoustic domain, humans group sounds alternating in pitch or amplitude as trochees, having initial stress, and sounds alternating in duration either as trochees or iambs (having final stress) depending on their native language (9-12). This study shows that zebra finches share the bias to group pitch alternations as trochees. However, they are unable to learn to discriminate between short strings of tones that alternated in duration.

Prosody perception. Every spoken sentence follows a certain prosodic pattern that is specific to a certain language. The prosodic patterns are created by changes in pitch, amplitude and duration of syllables. Infants can distinguish the prosodic patterns of different languages and generalize this to new syllables with the same prosody, and, as mentioned above, infants use prosody when segmenting speech (13-17). This shows that prosody is important to understand and segment speech and that the perception of prosodic information has an early onset in language learners.

In CHAPTER 4 we studied whether zebra finches are sensitive to the prosody in human speech and, if so, which cues they are attending to. The songs produced by zebra finches can vary in prosody-like features. For example, they can alter the amplitude of their song dependent on the distance of the receiver (18) and sing at different speeds for directed or undirected song (19). In our experiment, the zebra finches were trained to discriminate between xyxy and xxy structures with prosodic stress on the first (Xyxy) or the last (xxyY) syllable. They were then tested with quadruplets in which the stress pattern and structural patterns were interchanged (xyxY and Xxyy); with quadruplets in which the prosodic cues presented were reduced to only pitch, only duration, only amplitude or a combination of two, and with quadruplets consisting of new syllables. The results show that zebra finches are sensitive to the same prosodic cues as humans are. They learn to discriminate between syllable quadruplets with initial or final prosodic stress and keep discriminating based on the prosodic pattern when the structure is changed. Furthermore, test items in which only one of the three prosodic features -pitch, amplitude or duration- is present, are still correctly discriminated by the zebra finches. They also generalize the prosodic pattern to quadruplets with new syllables. This shows that a sensitivity to

prosody is not unique for humans or human language, but is shared with a songbird species (20). Prosody perception might be one of the cognitive mechanisms that did not evolve specifically for language.

Perception of syntactic patterns

When learning language, it is important to understand the general regularities that underlie the specific language. These could be the natural intonation pattern of a sentence, the rhythm of the speech or the grammar rule defining the order of words. Infants seem to learn these abstract patterns rather effortlessly during early language acquisition. However, this ability to abstract structural regularities has been difficult to demonstrate in non-human animals. Below I will discuss the two chapters of this thesis in which we have studied the pattern abstraction abilities of zebra finches and budgerigars.

Rhythmic patterns. Each musical piece has its own rhythm and each rhythm has its own beat pattern which remains recognizable when the song is played faster or slower. Infants are able to perceive these beat patterns and can group faster and slower versions of a beat pattern as belonging together. Moreover, newborns are already sensitive to a regular beat, and show increased brain activity when an expected beat is missing (21).

However, rhythm perception is not unique to humans. The study of rhythm perception in non-human animals was sparked by the discovery of a sulphur-crested cockatoo, Snowball, who was able to synchronize his movements to the beat of different songs (22, 23). Most impressively, he dynamically adjusted the tempo of his movements with changes in the tempo of the beat, indicating that he did respond to the abstract beat pattern. A few other bird species have shown to be able to discriminate between different beat patterns and generalize this to tempo transfers (starlings (24, 25), jackdaws (26), pigeons (27)). However, up to today it are mostly parrots that have shown impressive abilities to entrain their own behaviour to different rhythms (28).

To test whether rhythm perception is especially well developed in the parrot clade or shared between more species, we tested a parrot species, the budgerigar, and a songbird, the zebra finch, on their rhythm perception and generalization abilities, as described in CHAPTER 5. Individuals of both species were trained to discriminate between a regular and an irregular beat pattern in a string of tones. Afterwards they were tested with different slower and faster versions of the pattern, with the single

tones remaining intact or being shorter and longer than during training. Both species learn to discriminate between the regular and irregular beat pattern and both are able to generalize this to certain modifications of the pattern. However, individuals of both species also pay attention to the exact duration of a tone or the duration of pauses between consecutive tones. This shows that although zebra finches and budgerigars can generalize the regularity of a beat pattern to some extent, they are also paying attention to the more local features of the sounds.

The results of chapter 5 (29) and other studies on zebra finch rhythm perception (30, 31) were related to recent work on the rhythmicity in vocal production by zebra finches (32-34) in CHAPTER 6. The fact that zebra finches can both produce and perceive regular patterns and that the same brain regions are involved in both processes, make them an ideal species to study the interaction between rhythm perception and production in more depth.

Structural patterns. Structural patterns are the order in which arbitrary sounds are organized. For example, a very simple structural pattern is an XXY pattern where a repetition of one item is followed by a different item. This pattern is often tested against an XYX pattern, in which the first and third item are identical. Infants can discriminate between these and more complex structural patterns. Moreover, they generalize the structure to novel sounds and show a preference for the familiar or novel structure (35-37).

In order to determine whether these abilities are specific to humans, birds (38-44), rats (45, 46) and primates (47, 48) have been tested on their abilities to discriminate between these and other structural patterns. These studies show that many species are able to discriminate between strings with the same sound items that follow two different structures. However, none of these species show clear generalization of the underlying structure to sound items of other categories. For example, starlings were trained to discriminate between structural patterns created from warble and rattle song syllables following $(AB)^n$ or A^nB^n grammar pattern (42). The test items also consisted of warble and rattle song syllables, sounds that did not belong to novel categories. Even though the birds had not heard these particular syllables before, they could categorize them as being warbles and rattles. This means that they did not need to learn the abstract pattern, but could also memorize the order of the A and B categories.

Two studies with zebra finches trained the birds to discriminate either between ABAB and AABB structures or between XYX and XXY structures (38, 39). In both studies

the birds managed to discriminate between the training items. Nevertheless, in both studies most birds were not able to generalize the structure of the training sets to novel sounds belonging to different categories.

The study described in CHAPTER 7 of this thesis adds to these results by showing that budgerigars are able to learn the structural rule of XYX and XXY patterns. They can generalize the structural rule to new sounds, belonging to new categories, with the same structure. They thus convincingly show to generalize abstract structural patterns to novel sounds belonging to different categories. The zebra finches, on the other hand, pay attention to the ordinal positions of the elements of the training items, rather than to the underlying structure. When discriminating between training triplets organized in XYX and XXY structure, they memorize the position of single elements and used this whilst categorizing new combinations of the training elements. However, they do not generalize the underlying structure to novel sounds. This shows that the ability to learn abstract grammar rules is present in a parrot species and is thus not specific to humans nor to language. However, it seems to be a rare ability amongst animals.

Conclusions

The work presented in this thesis shows two different aspects of language perception: the perception of the phonological patterns of speech and the perception of abstract, syntactic patterns of language. The zebra finches are able to segment strings based on pause-cues, to perceptually group tones alternating in pitch and to perceive, learn and generalize the prosody of human speech. This shows that these perceptual abilities for phonological patterns are not unique to human language. Moreover, as we find similar abilities in other animal species (e.g. primates (49, 50), rats (51, 52), java sparrows (53) and budgerigars (54)), it is likely that these perceptual abilities were present pre-linguistically and may have formed a basis for the evolution of more complex sound structures and will have contributed to the evolution of speech as a way to express language. Learning abstract, syntactic patterns has proven to be much harder for both zebra finches and budgerigars. They show limited abilities to generalize regular beat patterns to tempo changes and only the budgerigars are able to generalize structural, grammatical patterns. These, more complex, cognitive mechanisms are also difficult to find in other non-human animals (see 55 for a review). The ability to learn abstract patterns is clearly present in humans, and some aspects of it might have evolved in parallel in a few species, but does not seem to be a broadly shared mechanism. This might also be due to the limited research that has been done so far, and the linguistic abilities of non-human animals might actually be less far from humans as is now assumed (56). However, the current state of the field suggests that the cognitive

mechanism for abstraction is likely to have evolved later than the abilities related to prosody and speech perception.

The ability to learn and generalize abstract patterns requires a range of perceptual abilities that seem to come natural to human infants. First, the infant needs to recognize that the sounds he or she is hearing are actually different from each other. Secondly, they need to recognize that these differences belong to one or more groups. Thirdly, they need to abstract the general pattern that defines a group and distinguishes the sounds in one group from those in another group. Lastly, the infants need to generalize these patterns to novel sounds following a different pattern. All these abilities seem to be present in language learning infants. The perceptual abilities of non-human animals show that the first two steps are readily shared between humans and other animals. The final two steps, however, might be more specific to humans and the language faculty.

In abstract pattern learning studies like the ones described in this thesis, discriminating the training stimuli could be based on both local and global information. Global information is the abstract pattern, like the grammar, the overall prosodic structure or the regularity of a string of sounds. Local information is, for example, the duration, pitch or ordinal position of a single sound item. These local features are sound specific and cannot be generalized to new sounds.

In humans, the use of these two types of information depends on the organization of the stimuli they are presented with (57, 58). Gerken (57) exposed two groups of infants to two sets of syllable triplets with an XXY structure. In one set the X and the Y element were a different syllable in each exemplar. In the second set, the X element differed but the Y element remained the same, always "di". Thus, both sets contained stable global information, the XXY pattern, but the second set also contained stable local information, the "di" syllable. Infants that heard the first set generalized the XXY structure to new exemplars, but infants that heard the second set only generalized to triplets where the final element was "di". This shows that, dependent on the information present in the stimuli, infants either use global or local information.

In our experiments there are both individual and species specific differences in the type of information used. In chapter 5 the regular and irregular stimuli could be discriminated based on their beat pattern. However, a correct discrimination could also be made by learning the exact position of the beats in the string. In both the zebra finches and the budgerigars, there were individuals that learned and generalized

the beat pattern of the strings, the global information. Nevertheless, there were also individuals that were unable to generalize the beat pattern and seemed to have learned the duration and pitch of each tone and their positions in the strings. In chapter 7 the birds had to discriminate between 5 XYX and 5 XXY triplets. The zebra finches learned the ordinal positions of the elements in the triplet, whereas the budgerigars learned the abstract pattern. This indicates that zebra finches and budgerigars might be able to use both local and global of information, but focus more on one type of information depending on the experiment, the species and the individual.

The species-specific difference in learning strategy, as seen in chapter 7, could be caused by the vocal-learning abilities, social auditory environment or differences in cognitive abilities between the two species. Zebra finches and budgerigars differ in the way they learn their vocalizations. Zebra finches are closed-end vocal learners, meaning that they have a sensitive phase in which they learn their songs and cannot learn new vocalizations after this phase. Budgerigars are open-end learners, they continue to learn new vocalizations throughout their lives. This also effects the amount of variability in their vocalizations. Where zebra finch males learn one particular song per individual, each budgerigar can produce a variety of complex songs (59, 60). These differences might have influenced the development of their auditory perception, causing them to focus on different types of information in our auditory stimuli. Dent et al. (61) showed that budgerigars are better than zebra finches at identifying complex acoustic stimuli. This might be an indication of different cognitive abilities related to acoustic perception. These differences in perceptual abilities might also be reflected in how well they can learn abstract patterns.

Comparative studies, like the ones described in this thesis, will continue to give us more insight in the cognitive mechanisms that underlie the evolution of music and language. Moreover, by examining both vocal learners and vocal non-learners, and by comparing closely related and more distantly related species with different cognitive abilities, we can better pinpoint the basic mechanisms that allowed for the development of language perception and production. Studies with primates, rats and birds already showed that the perception of phonological features is not related to the vocal learning capacity or to a having a relatively recent common ancestor with humans. These phonological features, changes in pitch, duration, amplitude or pause length, are also the sound features that seem to bring music and language together. Simplistically spoken, music and language are both strings of sounds with a pattern of modifications of prosody-like features. Studying how animals perceive these modification patterns and whether they are sensitive to particular musical or linguistic

patterns, sets the basis to find the roots of music and language and their possible interaction.

With this in mind, we have created a strong foundation to develop systematic studies to further explore the music and language capacities of animals. Seeing how much there is still to be unravelled and the speed with which the field develops, leaves me to think that we will continue to discover many more shared cognitive abilities between humans and non-humans. Although the human language capacity is unique in its kind, the mechanisms underlying it might all share their roots with other animal species.

REFERENCES

1. Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-monthold infants. Science 274:1926–28.

2. Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of memory and language, 44(4), 548-567.

3. Thiessen, E.D. & Saffran, J.R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. Developmental Psychology, 39, 706–716.

4. Bion, R.A., Benavides-Varela, S. & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. Language and Speech, 54(1), 123–140.

5. Hay, J.S.F. & Saffran, J.R. (2012). Rhythmic grouping biases constrain infant statistical learning. Infancy, 17(6), 610–641.

6. ten Cate, C. & Slater, P.J.B. (1991), Song learning in zebra finches—how are elements from 2 tutors integrated. Anim Behav 42:150–152.

7. Williams, H. & Staples, K. (1992). Syllable chunking in zebra finch (Taeniopygia guttate) song. J Comp Psychol 106(3):278–286.

8. Spierings, M., de Weger, A. & ten Cate, C. (2015). Pauses enhance chunk recognition in song element strings by zebra finches. Anim Cognit 18:867–874.

9. Hayes, B. (1985). Iambic and trochaic rhythm in stress rules. In Proceedings of the Thirteenth meeting of the Berkeley Linguistics Society, ed. by M. Niepokuj, et. al., 429–446. Berkeley, CA: Berkeley Linguistics Society.

10. Iversen, J.R., Patel, A.D. & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. The journal of the Acoustical Society of America 124(4), 2263-2271.

11. Yoshida, K.A., Iversen, J.R., Patel, A.D., Mazuka, R., Nito, H., Gervain, J., et al. (2010). The development of perceptual grouping biases in infancy: A Japanese–English cross-linguistic study. Cognition, 115, 356–361.

12. Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T. & Hohle, B. (2013). Native language affects rhythmic grouping of speech. The Journal of the Acoustical Society of America, 134(5), 3828–3843.

13. Jusczyk, P.W., Cutler, A. & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. Child Development, 64(3), 675–687.

14. Nazzi, T., Jusczyk, P.W. & Johnson, E.K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. Journal of memory and language, 43(1), 1-19.

15. Mehler, J. & Christophe, A. (1995). Maturation and learning of language during the first year of life. In M. S. Gazzaniga (Ed.), The Cognitive Neurosciences (pp. 943-954): Bradford Books / MIT Press.

16. Nazzi, T., Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. Journal of Experimental Psychology: Human Perception and Performance, 24(3), 756-766.

17. Ramus, F., Hauser, M.D., Miller, C., Morris, D. & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. Science, 288, 349– 351.

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

19. Sossinka, R. & Bohner, J. (1980). Song types in the zebra finch (Poephila guttata castanotis). Z. Tierpsychol. 53, 123 – 132.

20. Spierings, M.J. & ten Cate, C. (2014). Zebra finches are sensitive to the prosodic features of human speech. Proc R Soc B 281(1787):20140480.

21. Winkler, I., Haden, G., Ladinig, O., Sziller, I. & Honing, H. (2009). Newborn infants detect the beat in music. Proc. Natl. Acad. Sci. USA, 106, 2468–2471.

22. Patel, A. D., Iversen, J. R. & Bregman, M. R. (2009a). Experimental evidence for synchornizatino to a musical beat in a nonhuman animal. Curr. Biol. 19, 827–830.

23. Patel, A. D., Iversen, J. R., Bregman, M. R. & Schulz, I. (2009b). Studying synchronization to a musical beat in nonhuman animals. Ann. N. Y. Acad. Sci. 1169, 459–469.

24. Hulse, S. H., Humpal, J., and Cynx, J. (1984). Processing of rhythmic sound structures by birds. Ann. N.Y. Acad. Sci. 423, 407–419.

25. Hulse, S., Humpal, J., and Cynx, J. (1984). Discrimination and generalization of rhythmic and arrhythmic sound patterns by European starlings (Sturnus vulgaris). Music Percept. 1, 442–464.

26. Reinert, J. (1965). Takt- und rhythmusunterscheidung bei dohlen. Z. Tierpsychol. 22, 623–671.

27. Hagmann, C. E. & Cook, R. G. (2010). Testing meter, rhythm, and tempo discriminations in pigeons. Behav. Process. 85, 99–110.

28. Schachner, A., Brady, T. F., Pepperberg, I. M. & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. Curr. Biol. 19, 831–836.

29. ten Cate, C., Spierings, M., Hubert, J. & Honing, H. (2016). Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species. Frontiers in Psychology, 7, 730.

30. Nagel, K.I., McLendon, H.M. & Doupe, A.J. (2010). Differential influence of frequency, timing, and intensity cues in a complex acoustic categorization task. Journal of Neurophysiology, 104(3), 1426-1437.

31. van der Aa, J., Honing, H. & ten Cate, C. (2015). The perception of regularity in an isochronous stimulus in zebra finches (Taeniopygia guttata) and humans. Behav. Processes eds D. Deutsch (London, UK: Elsevier/Academic Press) 115, 37–45.

32. Norton, P. & Scharff, C. (2016). 'Bird song metronomics': Isochronous organization of zebra finch song rhythm. Front. Neurosci. 10:309.

33. Benichov, J.I., Benezra, S.E., Vallentin, D., Globerson, E., Long, M.A. & Tchernichovski O. (2016a). The forebrain song system mediates predictive call timing in female and male zebra finches. Curr. Biol. 26, 309–318.

34. Benichov, J.L., Globerson, E. & Tchernichovski, O. (2016b). Finding the beat: From socially coordinated vocalizations in songbirds to rhythmic entrainment in humans. Front. Hum. Neurosci. 10:255.

35. Marcus, G.F., Vijayan, S., Rao, S.B. & Vishton, P.M. (1999). Rule learning by seven-month-old infants. Science 283(5398):77-80.

36. Gomez, R.L. & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70:109–35.

37. Christiansen, M.H., Conway, C.M. & Curtin, S. (2000). A connectionist single mechanism account of rule-like behavior in infancy. In: Proceedings of the 22nd Annual Conference of the Cognitive Science Society. Erlbaum.

38. Chen, J., van Rossum, D. & ten Cate, C. (2014). Artificial grammar learning in zebra finches and human adults: XYX vs XXY. Anim Cogn 18:151–164. doi: 10.1007/s10071-014-0786-4

39. van Heijningen, C.A., de Visser, J., Zuidema, W. & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. PNAS USA 106:20538–20543.

40. van Heijningen, C.A., Chen, J., van Laatum, I., van der Hulst, B. & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? Anim Cogn 16:165–175.

41. Stobbe, N., Westphal-Fitch, G., Aust, U. & Fitch, W.T. (2012). Visual artificial grammar learning: Comparative research on humans, kea (Nestor notabilis) and pigeons (Columba livia). Philos Trans R Soc Lond B Biol Sci 367(1598): 1995–2006.

42. Gentner, T.Q., Fenn, K.M., Margoliash, D. & Nusbaum, H.C. (2006). Recursive syntactic pattern learning by songbirds. Nature 440(7088):1204–1207.

43. Comins, J.A. & Gentner, T.Q. (2014). Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. Anim Cogn 17(5):1023–1030.

44. Comins, J.A. & Gentner, T.Q. (2013). Perceptual categories enable pattern generalization in songbirds. Cognition 128(2):113–118.

45. Murphy, R.A., Mondragón, E. & Murphy, V.A. (2008). Rule learning by rats. Science 319(5871): 1849–1851.

46. de la Mora, D.M. & Toro, J.M. (2013). Rule learning over consonants and vowels in a non-human animal. Cognition 126(2):307–312.

47. Endress, A.D., Carden, S., Versace, E. & Hauser, M.D. (2010). The apes' edge: Positional learning in chimpanzees and humans. Anim Cogn 13(3):483–495.

48. Fitch, W.T. & Hauser, M.D. (2004). Computational constraints on syntactic processing in a nonhuman primate. Science 303(5656):377–380.

49. Ramus, F., Hauser, M.D., Miller, C., Morris, D. & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. Science, 288, 349– 351.

50. Tincoff, R., Hauser, M., Tsao, F., Spaepen, G., Ramus, F. & Mehler, J. (2005). The role of speech rhythm in language discrimination: further tests with a non-human primate. Dev. Sci. 8, 26-35.

51. Toro, J.M., Trobalon, J.B. & Sebastian-Galles, N. (2003). The use of prosodic cues in language discrimination tasks by rats. Anim. Cogn. 6, 131 – 136.

52. Toro, J.M., Trobalon, J.B. & Sebastian-Galles, N. (2005). Effects of backward speech and speaker
variability in language discrimination by rats. J. Exp. Psychol. Anim. B 31, 95 – 100.

53. Naoi, N., Watanabe, S., Maekawa, K. & Hibiya, J. (2012). Prosody discrimination by songbirds (Padda oryzivora). PLoS ONE 7, e47446.

54. Hoeschele, M. & Fitch, T. (2016). Phonological perception by birds: budgerigars can perceive lexical stress. Animal Cognition, 19(3), 643-654.

55. ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. 367, 1984–1994.

56. ten Cate, C. (2016). Assessing the uniqueness of language: Animal grammatical abilities take center stage. Psychon Bull Rev.

57. Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. Cognition 98(3):B67–B74.

58. Kovacs, A.M. (2014). Extracting regularities from noise: Do infants encode patterns based on same and different relations? Lang Learn 64(2):65–85.

59. Farabaugh, S.M., Brown, E.D. & Dooling, R.J. (1992). Analysis of the warble song of the budgerigar (Melopsittacus undulates). Bioacoustics 4(2):111–130.

60. Tu, H.W., Smith, E.W. & Dooling, R.J. (2011). Acoustic and perceptual categories of vocal elements in the warble song of budgerigars (Melopsittacus undulatus). J Comp Psychol 125(4):420–430

61. Dent, M.L., Welch, T.E., McClaine, E.M. & Shinn-Cunningham, B.G. (2008). Species differences in the identification of acoustic stimuli by birds. Behav Processes 77(2):184–190.

ADDENDUM



NEDERLANDSE SAMENVATTING

Taal is uniek voor mensen. Hoewel alle dieren communiceren, is daarin nog niet dezelfde complexiteit en structuur gevonden als in menselijke taal. Vanwege zijn uniekheid wordt er ook hevig gediscussieerd over het ontstaan en de evolutie van taal. Hoe communiceerden onze verre voorouders? Wat waren de eerste klanken en structuren, en hoe heeft evolutie hierop kunnen doorbouwen? In 1866 liepen de debatten hierover zo op, dat de Société de Linguistique de Paris een ban legde op de discussie over de evolutie van taal. Er was te weinig empirisch bewijs om alle verschillende hypothesen te onderbouwen.

Tegenwoordig wordt er al een aantal decennia weer onderzoek gedaan naar de evolutie van taal. Omdat er geen fossielen van gesproken taal bestaan, moet er op een andere manier onderzoek gedaan worden. Een recente, maar belangrijke, manier om empirisch onderzoek te doen naar taalevolutie, is met vergelijkend onderzoek. Hiermee vergelijken onderzoekers de cognitieve eigenschappen van mensen en andere diersoorten. Eigenschappen die gedeeld zijn tussen mensen en andere dieren, zijn de eigenschappen die waarschijnlijk al vroeg in de evolutie ontstaan zijn. Deze eigenschappen kunnen aanwezig zijn geweest in onze verre voorouders en de basis hebben gelegd waarop de evolutie van taal voortgeborduurd kan hebben.

Gedurende mijn promotie ben ik onderdeel geweest van een projectgroep waarin we samen vergelijkend onderzoek deden. Deze projectgroep bestond uit wetenschappers die onderzoek deden naar taalontwikkeling bij baby's en wetenschappers die computermodellen maakten voor de betrokken cognitieve processen. Met deze interdisciplinaire groep keken we naar verschillen en overeenkomsten tussen de taalgerelateerde cognitieve eigenschappen jonge mensen en vogels. Daarnaast onderzochten we welke aspecten van taal belangrijk zijn voor mensen en welke eigenschappen computermodellen moeten hebben om het leerproces te kunnen simuleren.

In mijn deel van dit project heb ik onderzoek verricht naar cognitieve vaardigheden gerelateerd aan taalperceptie bij twee vogelsoorten, zebravinken (zangvogels) en parkieten (papegaaien). Deze twee dieren staan fylogenetisch ver van de mens af, maar delen één belangrijke eigenschap met mensen: vocaal leren. Zowel zangvogels als papegaaien hebben vocale voorbeelden nodig om zelf hun zang te leren. Dit is vergelijkbaar met hoe mensen taal moeten horen om taal te kunnen produceren.

Andere primaten, zoals chimpansees en orang-oetangs, zijn geen vocale leerders. Zij produceren dezelfde geluiden, ongeacht of ze een vocaal voorbeeld hebben gehad.

Het werk in dit proefschrift bestaat uit twee delen. Ten eerste heb ik onderzoek gedaan naar eigenschappen gerelateerd aan de perceptie van klankstructuren. Hierbij heb ik bijvoorbeeld onderzocht hoe zebravinken lange reeksen geluiden in korte stukken opdelen en of ze gevoelig zijn voor de intonatie van menselijke spraak. In het tweede deel heb ik de complexere eigenschappen onderzocht die gerelateerd zijn aan de perceptie van taal. Hierin heb ik onderzocht of zebravinken en parkieten abstracte structuren kunnen herkennen in geluid, zoals het algemene ritme van een reeks tonen of de grammaticale structuur van een aantal geluidselementen.

Perceptie van klankstructuren

In hoofdstuk 2 werd onderzocht hoe zebravinken lange reeksen zebravink-geluiden in kortere delen konden opdelen. Ze konden hiervoor gebruik maken van de mate waarin twee klanken na elkaar voorkwamen of, in een tweede groep dieren, van nauwelijks waarneembare stiltes die tussen de delen ingevoegd waren. Mensen kunnen dit allebei vanaf jonge leeftijd al gebruiken om klankreeksen op te delen De resultaten lieten zien dat zebravinken wel deze korte stiltes gebruikten, maar niet welke klanken samen voorkomen.

In hoofdstuk 3 werd ook naar lange reeksen gekeken, maar nu van tonen. Deze tonen waren afwisselend hoog-laag in een groep of afwisselend lang-kort in de andere groep. Mensen groeperen dit soort afwisselingen automatisch als groepen met een hoog-laag of kort-lang patroon. De zebravinken groepeerden de afwisselingen in toonhoogte op een vergelijkbare manier met mensen. Ze konden niet leren om onderscheid te maken in tonen die wisselden in toonduur.

In hoofdstuk 4 is de perceptie van prosodische patronen onderzocht in zebravinken. Prosodische patronen zijn veranderingen in spraak door het veranderen van de frequentie, duur en amplitude van de lettergreep. Door veranderingen in prosodie kan er een andere betekenis aan dezelfde zin gegeven worden. De zebravinken in deze studie konden goed leren om onderscheid te maken tussen twee verschillende prosodische patronen. Ook als er minder prosodische veranderingen aanwezig waren, maakten ze nog goed onderscheid en zelfs als de spraakklanken veranderden, konden de zebravinken nog het prosodische patroon volgen.

Deze drie studies tonen aan dat de perceptie van prosodische structuren in klanken niet uniek is voor mensen en taal, maar gedeeld is met een zangvogel.

Perceptie van abstracte structuren in geluid

Voor het begrijpen van taal is het cruciaal dat men de onderliggende structuur van een reeks geluiden kan herkennen. De voorgaande hoofdstukken onderzochten vooral of verschillende klanken onderscheiden en gegroepeerd konden worden. In deze hoofdstukken onderzochten we of vogels de abstracte structuren van een reeks geluiden kunnen leren en generaliseren.

In hoofdstuk 5 werd onderzocht of zebravinken en parkieten onderscheid kunnen maken tussen regelmatige en onregelmatige toonreeksen. We testen ze met versnelde en vertraagde versies van de trainingsgeluiden om te zien of ze echt de regelmatigheid hadden geleerd, of dat ze bijvoorbeeld de absolute afstand tussen twee tonen hadden geleerd. De resultaten lieten zien dat beide soorten wel iets van de algemene regelmaat hadden geleerd, maar dat ze ook gelet hadden op details, zoals de precieze lengte van een toon en de lengte van de pauze tussen twee tonen in. Dit onderzoek liet variatie tussen individuen zien, maar toonde ook aan dat generalisatie van abstracte structuren voor zebravinken en parkieten niet vanzelfsprekend is.

In hoofdstuk 6 werd dit type onderzoek gerelateerd aan onderzoek naar de structuur in de geluiden die zebravinken produceren. Zowel de zang als andere vocale interacties van zebravinken volgen een bepaalde regelmaat. De combinatie van perceptie en productie maakt zebravinken een goed diersoort voor verder onderzoek naar ritme en regelmaat.

Hoofdstuk 7 onderzocht ook of zebravinken en parkieten een abstracte structuren konden herkennen, maar dit keer was het een grammaticale structuur. Ze leerden om onderscheid te maken tussen geluiden in een XYX structuur en geluiden in een XXY structuur. Daarna werden ze getest met nieuwe geluiden die dezelfde structuren volgden. De zebravinken leerden de exacte positie van elk geluid in de training. De parkieten leerden wel de abstracte structuur. Dat betekent dat ze totaal nieuwe geluiden met een XYX structuur in dezelfde categorie plaatsten als de trainingsgeluiden met een XYX structuur. Dit is een van de eerste duidelijk bewijzen dat dieren abstracte structuren kunnen leren en herkennen.

De resultaten van deze en eerdere studies laten zien dat de perceptie van klankstructuren gedeeld is met meerdere soorten. Niet alleen vogels, maar ook ratten

en apen kunnen verschillen en structuren in klanken herkennen. De perceptie van de abstracte structuren van reeksen klanken is daarentegen veel lastiger. Zowel de zebravinken als de parkieten hadden moeite met het leren van het abstracte ritmische structuur van reeksen toontjes. De grammaticale structuren konden wel door de parkieten geleerd worden, maar zij behoren hiermee tot de weinige soorten die dit kunnen.

De studies in dit proefschrift laten zien dat vergelijkend onderzoek inzicht kan geven in de cognitieve eigenschappen die aan de basis van taalevolutie kunnen hebben gelegen. Met deze eerste stappen kunnen we verder ontdekken welke onderdelen van taal perceptie wel, en welke niet, aanwezig zijn in verschillende diersoorten. Zo zullen we de volgende puzzelstukjes vinden om steeds meer te ontdekken over het ontstaan van onze taal.

ACKNOWLEDGEMENTS

First of all, of course, I would like to thank Carel ten Cate. You have been a great supervisor and a well needed balance. You have guided me through this PhD track with patience and cheerfulness which I really appreciate.

Thanks to my committee members, I am grateful that you are willing to read and comment on my thesis.

Many thanks to the project group, Claartje Levelt, Jelle Zuidema, Raquel G. Alhama and Andreea Geambasu. In our monthly meetings we could discuss experimental designs, results, papers and conferences. Doing a PhD in such a creative and helpful group made the experience much better. Raquel and Andreea, it was great starting our PhD together and working on similar experiments. I loved working together, not only to talk about the work, but also about all the other aspects of doing a PhD.

Thanks to Elisabetta Versace and Giorgio Vallortigara for collaborating with me and hosting me in Italy. With your knowledge, experience and enthusiastic work regime, you created a great place for me to work. Thank you Elisabetta for making me feel at home and for your help with all the things that required Italian. I am glad we are still working together.

Thank you Henkjan Honing for our fruitful collaboration. Your unlimited enthusiasm is inspiring and contagious. I hope we can continue to collaborate in the future.

Thanks to all the people of the behavioural biology lab in Leiden. You have created a comfortable and stimulating environment to work in, with so many friendly and helpful colleagues. I really enjoyed being part of this group and will definitely miss you.

My thanks also goes out to all the students and animal caretakers who have helped me during my PhD. With your (hard) work you have helped me with the experiments, added interesting thoughts to our discussions and have created an interesting and renewing work environment. Jeroen, thank you for taking over so well when I was in Italy.

Thank you Andreea and Erwin for being my paranimphs. You both know all about the work in this thesis and gave me new ideas from your fresh view on the project. Thank you for never getting bored with my bird stories.

Many thanks to my family, my parents who seem to have become the worlds' most enthusiastic bird-research followers, and my sisters and brothers who never fail to celebrate any achievement. And thanks to all of my dear friends, you make life a lot nicer and happier.

Last, but not least, thanks to Robert. Thank you for being enthusiastic about everything, for seeing problems as challenges, and for finding great new opportunities when plans change.

CURRICULUM VITAE

Profile

Date of birth	18-07-1987
Place of birth	Voorburg, Netherlands
Scientific interests	Animal cognition, Auditory production & perception, Comparative cognition

Education

1999-2005 VWO, St. Gregorius College, Utrecht

2006-2009 BSc Biology, Leiden University

Internship at Behavioural Biology, Leiden University.
Under supervision of dr. Katharina Riebel I studied the social learning strategies of zebra finches from different brood sizes. This work resulted in a publication in *Animal Behaviour*.

2009-2011 MSc Neuroscience and Cognition, Utrecht University

Internship at Behavioural Biology, Utrecht University
Under supervision of dr. Simon Reader I studied whether guppies could learn different foraging strategies dependent on the context.

Internship at Experimental Psychology, Cambridge University
During this internship I studied theory of mind in rooks, a corvid species under supervision of prof. Nicola Clayton.

Honours track in entrepreneurship
During this honours track we learned to set up innovative projects and how to develop as an entrepreneur.

2012-2016 PhD, Behavioural Biology, IBL, Leiden University

Under supervision of prof. dr. Carel ten Cate I studied cognitive abilities related to language perception in two bird species, zebra finches and budgerigars. This work was conducted as part of an interdisciplinary project group with developmental linguists and

computational linguists. All zebra finch and budgerigar studies were conducted at Leiden University. Four months of the project were spent at the University of Trento to work together with prof. Giorgio Vallortigara and dr. Elisabetta Versace on perception of chicks. During my PhD I supervised several BSc and MSc students during their internships and I presented my worked at several national and international conferences. Furthermore, I was closely involved in the designing, realization and implementation of new skinnerboxes and a new software program. This PhD project was funded by NWO-GW grant 360.70.452.

Grants and Awards

Hurford prize, Evolang, New Orleans, 2016
Best presentation at Evolang, an international conference on the evolution of language.

Famelab National Final, 2015
Famelab is an international science communication competition, in which young scientists explain their research to a laymen's audience in just three minutes.

Presentation award at the NVG conference, Netherlands, 2015
Best presentation at NVG, a national conference on animal behaviour.

LUF Travel Grant
Leiden University travel grant to visit the Behaviour conference in Cairns, Australia, 2015

Outreach

During my PhD I have presented my work on several public events in and outside of the Netherlands. Furthermore, I have explained my work on national television, national and international radio stations and in several newspapers.

PUBLICATIONS

Geambasu, A.*, **Spierings, M.J.***, ten Cate, C. & Levelt, C.C. (in preparation) Tell me what to learn: The effects of task-specific variables on artificial grammar learning and generalization. * authors share first authorship

Versace, E.*, **Spierings, M.J.***, ten Cate, C. & Vallortigara, G. (in preparation) Imprinting generalization of patterns in young domestic chicks. * authors share first authorship

Spierings, M.J., Hubert, J. & ten Cate, C. (under review) Selective auditory grouping by zebra finches: testing the Iambic-Trochaic law.

Spierings, M.J. & ten Cate, C. (2016) Zebra finches as a model species to understand the roots of rhythm. *Frontiers in Neuroscience*, 10, 345.

Spierings, M.J. & ten Cate, C. (2016) Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences (PNAS)*, 113(27), E3977-E3984.

ten Cate, C. *, **Spierings, M.J. ***, Hubert, J. & Honing, H. (2016) Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species. *Frontiers in psychology*, 7:730. * authors share first authorship

Kriengwatana, B., **Spierings, M.J.** & ten Cate, C. (2016) Auditory discrimination learning in zebra finches: Effects of sex, early life conditions and stimulus characteristics. *Animal Behaviour*, 116, 99-112.

Spierings, M.J., de Weger, A. & ten Cate, C. (2015) Pauses enhance chunck recognition in song element strings by zebra finches. *Animal cognition*, 18(4), 867-874.

Spierings, M.J. & ten Cate, C. (2014) Zebra finches are sensitive to prosodic features of human speech. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1787), 20140480.

Riebel, K., **Spierings, M.J.**, Holveck, M-J. & Verhulst, S. (2012) Phenotypic plasticity of avian social-learning strategies. *Animal Behaviour*, 84(6), 1533-1539.

ten Cate, C., Bruins, W.S., den Ouden, J., Egberts, T., Neevel, H., **Spierings, M.J.**, van der Burg, K. & Brokerhof, A.W. (2009) Tinbergen revisited: a replication and extension of experiments on the beak colour preferences of herring gull chicks. *Animal Behaviour*, 77(4), 795-802.

Conference proceedings

Spierings, M.J. & ten Cate, C. (2016) Rule Learning In Birds: Zebra Finches Generalize By Positional Similarities, Budgerigars By The Structural Rules. In S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér & T. Verhoef (eds.) *The Evolution of Language: Proceedings of the 11th International Conference* (EVOLANG11).

Geambasu, A., **Spierings, M.J.**, ten Cate, C. & Levelt, C.C. (2016) Effects Of Task-specific Variables On Auditory Artificial Grammar Learning And Generalization. In S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér & T. Verhoef (eds.) *The Evolution of Language: Proceedings of the 11th International Conference* (EVOLANG11).

Spierings, M.J. & ten Cate, C. (2014) Element string segmentation by zebra finches. In Lišková, S. & Varadínová, Z. (Eds.) *The 7th European Conference on Behavioural Biology 2014*.

Spierings, M.J. & ten Cate, C. (2014) Prosodic cue weighting by zebra finches. In Cartmill, E. A., Roberts, S. G., Lyn, H., & Cornish, H. (Eds.) *The Evolution of Language: Proceedings of the 10th International Conference* (EVOLANG10).

Geambasu, A., Levelt, C.C., **Spierings, M.J.** & ten Cate, C. (2014) Artificial grammar learning in infants, adults, and songbirds: What is shared, what is learned? In Cartmill, E. A., Roberts, S. G., Lyn, H., & Cornish, H. (Eds.). *The Evolution of Language: Proceedings of the 10th International Conference* (EVOLANG10).

