



Universiteit
Leiden
The Netherlands

Linguistic birds : exploring cognitive abilities in zebra finches by using artificial grammars

Chen, J.

Citation

Chen, J. (2014, October 16). *Linguistic birds : exploring cognitive abilities in zebra finches by using artificial grammars*. Retrieved from <https://hdl.handle.net/1887/29164>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/29164>

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

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/29164> holds various files of this Leiden University dissertation.

Author: Chen, Jiani

Title: Linguistic birds : exploring cognitive abilities in zebra finches by using artificial grammars

Issue Date: 2014-10-16

Linguistic Birds

Exploring cognitive abilities in zebra finches by
using artificial grammars

Jiani Chen

Chen, Jiani
Linguistic birds
Exploring cognitive abilities in zebra finches by using artificial grammars
Dissertation Leiden University
ISBN: 978-94-6203-671-0
An electronic version of this thesis is available at: https://openaccess.leidenuniv.nl
Cover design by Shuyun Chen (陈淑芸)
Printed by CPI-Koninklijke Wöhrmann, Zutphen

Linguistic Birds

Exploring cognitive abilities in zebra finches by using
artificial grammars

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 voor Promoties
te verdedigen op donderdag 16 oktober 2014
klokke 10:00 uur

door

Jiani Chen

geboren te Guangdong, China
in 1984

Promotiecommissie

Promotor:

Prof. Dr. Carel J. ten Cate

Overige leden:

Dr. Juan Manuel Toro (Universitat Pompeu Fabra, Barcelona)

Dr. Willem H. Zuidema (Universiteit van Amsterdam)

Prof. Dr. Clara C. Levelt

Prof. Dr. Herman P. Spaink

This work was supported by the China Scholarship Council (CSC).

Contents

Chapter 1	General Introduction	1
Chapter 2	Zebra finches can use positional and transitional cues to distinguish vocal element strings	11
Chapter 3	Zebra finch are able to learn prefixation and suffixation patterns	27
Chapter 4	Artificial grammar learning in zebra finches and human adults: XYX vs XXY.	41
Chapter 5	Bridging the gap: learning of non-adjacent dependencies in a songbird	67
Chapter 6	Summary and conclusion	89
	Dutch summary	95
	Acknowledgements	97
	Curriculum Vitae	98
	Publications	99

Chapter 1

General Introduction

Among animals, all kinds of senses are used for communication. There are chemical communication systems, systems mediated by behavioral displays (e.g. facial expressions and gestures), vocal communication systems, and so on. Among all communication systems, human language is the most complex one known to date. Understanding the evolution of the language faculty is one of the greatest scientific challenges; one that has given and still gives rise to hot debates (Hauser, Chomsky et al., 2002; Hauser, Yang et al., 2014; Jackendoff & Pinker, 2005). The origin of language is difficult to track since there are neither fossils nor clearly identifiable, biologically heritable entities that directly link to such behavior. Comparative studies exploring the linguistic capacities in nonhuman animals provide an alternative perspective for understanding the evolution of language, while at the same time revealing more about the cognitive capacities of animals. Although the specific functioning of the communication system of each species is unique, studying what is and is not shared between different species can help us to identify what might have been at the evolutionary basis for a complex trait like human language or what might have been the selection pressures that have led to different evolutionary trajectories.

The aim of this thesis is to shed light on whether some capacities that are considered linked to, or characteristic for, language are shared between humans and nonhuman animals, which may help to understand the basic cognitive abilities from which the evolution of human language may have arisen. I start with comparing human language with other communication systems to illustrate how distinct language is and to make clear that, in order to understand the cognitive mechanisms that might have been at the basis for the evolution of the language faculty, one should not restrict oneself to examining the communication systems of other animals. I suggest that a possibly better way to study the evolutionary basis of language is to study the broader cognitive abilities of nonhuman animals and map these abilities to human linguistic capacities to check which are and which are not shared. This is the core theme of this thesis.

Human language and other communication systems

Language is highly structured and hierarchically organized. Language users must therefore possess a certain level of cognitive skills or capacities that allow them to comprehend language by decomposing the hierarchically structured units. Over the years, the question which aspects of human language are also found in non-human animal communication systems has been hotly debated (see Hauser, Yang et al., 2014 for a recent example). Humans have tried for decades to teach language to our closest primate relatives, but without much success (Bolhuis & Wynne, 2009). This indicates that some components that are important to language may not be available to our relatives. One key prerequisite tied to language is vocal imitation (Fitch, 2000), which enables learners to obtain the correct sounds and structures of their language from other individuals. Our primate cousins show almost no evidence of vocal imitation (Whiten & Ham, 1992). For example, monkeys produce basically all their species-typical call types even when they were reared in isolated conditions (Arbib, Liebal et al., 2008). A cross-fostering study showed that rhesus and Japanese macaques cross-fostered to each other produced limited changes in their species-specific vocalizations when compared with the

control individuals that were raised by conspecifics (Owren, Dieter et al., 1993). Also, few studies show vocal dialects in primates, suggesting the lack of vocal imitation among nonhuman primates (Hauser, Chomsky et al., 2002).

Different from our primate cousins, distantly related species such as some marine mammals, bats and songbirds were found to have greater vocal learning abilities (Boughman, 1998; Janik & Slater, 1997; Janik & Slater, 1998; Knörnschild, Nagy et al., 2010; Wilbrecht & Nottebohm, 2003). Vocal learning in songbirds is especially well studied and striking parallels have been found between song learning and speech acquisition. Songbirds and humans both learn their vocalizations in early life, and the sensory-motor systems for vocal acquisition in songbirds and humans share notable similarities (Bolhuis, Okanoya et al., 2010; Doupe & Kuhl, 1999). Both human language and birdsong are also hierarchically organized, following certain syntactic constraints. However, the nature of the syntactic constraints is generally considered as the core difference between language and any non-human animal communication system known to date (Hauser, Chomsky et al., 2002). Unlike human language, bird songs can be considered as a regular language (characterized by finite-state machines, which are computational models that consist of a finite number of states and the transition functions between these states). In contrast, human languages demonstrate the ability of recursion, which is the possibility to infinitely embed syntactical structures inside an existing structure, and consequently has the features of a context-free language, allowing an infinite range of expressions (Berwick, Okanoya et al., 2011; Hauser, Chomsky et al., 2002). Recursive constructs such as $a^n b^n$ (the bird the cat chased flew, a_1 = the bird, a_2 = the cat, b_2 = chased and b_1 = flew) occur in many human languages and rely on nested dependencies. To detect whether an animal can produce a context-free pattern, one needs to demonstrate that the animal can correctly pair the $a(s)$ and the $b(s)$ (Berwick, Okanoya et al., 2011). Dependency, especially the nonadjacent ones (such as illustrated in the sentence above, in which a_1 and b_1 have a nonadjacent dependency involving event knowledge) is suggested to be important and critical characteristic of human language (Everaert & Huybregts, 2013), and no evidence to date has shown the existence of nonadjacent dependencies in the communication system of any other animal.

Altogether, the comparison between communication in humans and nonhumans tells us that human language is highly distinctive, and that the spontaneous usage of some features in human language (such as recursion and nonadjacent dependency) are not found in the communication systems of nonhuman animals. But does this mean that the linguistic mechanisms underlying these features are as distinctive as one might think? In other words, is it possible that these features, although not shown in non-human animal communication, can still be recognized by domain-general cognitive mechanisms not specific to language? It is possible that the cognitive mechanisms that allow the detection of these features evolved for different purposes in nonhuman animals and thus are not obvious in their communication systems. This idea is supported by the presence of some advanced cognitive capacities that are not reflected in the communicative content of nonhuman animals. For example, animals are able to anticipate future (Correia, Dickinson et al., 2007; McKenzie, Cherman et al., 2004; Mulcahy & Call, 2006; Raby, Alexis et al., 2007), but there is no evidence showing that this

capacity is reflected in their communications (they do not “talk” about things and events that are not here and now). Evidence of such abilities that are not reflected in communication also comes from studies of artificial language learning. Animals are able to perceive and generalize some types of grammatical structures not present in their communication signals (Ravignani, Sonnweber et al., 2013; Saffran, Hauser et al., 2008).

What to compare

Some capacities required for language learning and comprehension may have evolved from domain-general cognitive mechanisms (Endress, Cahill et al., 2009; Endress, Carden et al., 2010; Endress, Dehaene-Lambertz et al., 2007; Endress, Nespor et al., 2009). Therefore, comparing human language with communication systems of other animals may not be the best way to get insight in the type of mechanisms that may have shaped the evolution of language. A better approach might therefore be to compare the presence and nature of cognitive capacities that are relevant to language learning. These cognitive capacities refer to what an animal can learn and not to what can be seen in the communication system of this animal. A capacity may not be ‘expressed’ in certain domains, but may be present in others. If such capacity is found to be shared by different species, it may have been at the basis of the evolution of the higher complexity as present in the structure of language. Therefore, study of shared capacities, no matter whether they are present in the communication systems of other animals or not, might help us to identify which underlying mechanisms are homologous, which are convergent and which are unique to different species.

This thesis reflects the line of thought presented above by exploring the presence of several cognitive abilities in a non-human animal, the zebra finch. Considering that the most important and possibly most unique (Hauser, Chomsky et al., 2002), feature of language is the computational mechanism that processes syntactical structures, this thesis concentrates on the nature of several learning abilities underlying linguistic structures, in particular some that currently attract considerable attention– namely, sequential learning and rule learning. The study species used to address the topic is the zebra finch, a model species for comparative studies on vocal learning. Birdsong and human speech share notable similarities (Berwick, Okanoya et al., 2011; Wilbrecht & Nottebohm, 2003) and there are striking parallels between the ways in which infants and young birds learn their vocalizations (Doupe & Kuhl, 1999). Zebra finches, as a model species, have been studied extensively, and notable similarities between these birds and humans have been found in studies in various areas. Zebra finches show functional similarities with human auditory and motor cortices in their pallial ‘song’ regions (Jarvis, Güntürkün et al., 2005). Studies on the molecular genetics of vocal learning showed that the forkhead box protein P2 (FOXP2) gene, which is crucial to language acquisition, also plays a key role in song learning in zebra finches (Fisher & Scharff, 2009). Moreover, behavioral studies suggested that zebra finches shared some similarities with humans in using cues for perceiving speech sounds (Ohms, Escudero et al., 2012; Ohms, Gill et al., 2010; Spierings & ten Cate, 2014). Such similarities between humans and zebra finches make these birds of special interest for examining cognitive capacities that link to language processing.

Thesis outline

This thesis consists of the present introduction and 4 chapters consisting of empirical studies, addressing questions from the domains of sequential learning and rule generalization.

Sequential learning

Components of communication systems, such as behavioral displays and sounds, are arranged in sequences. Also verbal language, the most complex system of communication, consists of sequences of speech sounds. The capacity to learn and encode sequences is basic to language learning and is expected to be constrained by sequential memory. A sequence can be memorized by using the positional information of the items. That is, encoding the items with their positions in the sequence (Henson, 1998). Phenomena like serial order intrusion errors suggest such a positioned encoding mechanism, which is supported by evidence from serial recall experiments (Endress, Cahill et al., 2009; Endress, Carden et al., 2010; Hitch, Fastame et al., 2005). Positional encoding also gained considerable support from studies in nonhuman animals. Studies of serial recall in primates and birds suggested that nonhuman animals rely mostly on positional information when memorizing sequences (Comins & Gentner, 2010; Endress, Carden et al., 2010; Orlov, Amit et al., 2006; Orlov, Yakovlev et al., 2000; Terrace, Son et al., 2003). However, studies on statistical learning showed that humans and monkeys were also sensitive to associations between items (Aslin, Saffran et al., 1998; Hauser, Newport et al., 2001; Kelly & Martin, 1994; Saffran, Aslin et al., 1996; Saffran, Johnson et al., 1999). In these studies, learners were good at using chaining-like mechanisms based on transitional probabilities between subsequent items in a string. This suggested that transitional information between the items may also play an important role in sequential learning. So far, the only study explicitly testing how birds (starlings) process sequential information (Comins & Gentner, 2010) suggested that they relied heavily on positional information, raising the question whether songbirds can also use transitional information. In **chapter 2** we examine this question by conducting an experiment with artificial language stimuli containing both positional and transitional information.

Rule generalization

Language is a rule- or structure-mediated system, therefore rule abstraction is considered to be a hallmark of human linguistic abilities (Hauser, Chomsky et al., 2002; Marcus, Vijayan et al., 1999; Pinker, 1991). Rule learning abilities of animals have been tested in various species and findings from these tests suggest that some rule learning capacities are shared between human and other animals. The subsequent chapters of this thesis address rule learning abilities that have so far not, or hardly, been explored for non-human animals.

Edge-based rule learning

Endress, Nespore et al. (2009) proposed that rule abstraction involving language learning might have been complemented and constrained by a toolbox of ‘perceptual or memory primitives’. One of these primitives is related to edge-based positional regularities, in which rule generalizations are constrained by the edge positions. Edge-based positional

regularities such as English regular past tense (-ed added to words, such as learned, recognized etc.) provides a strong case for establishing the view of mental computation as rule-based manipulation of symbol systems (Marslen-Wilson & Tyler, 1998). Tracking the edges of a sequence is not a phenomenon only found in humans. Evidence from a non-human primate study (Endress, Cahill et al., 2009) suggested an evolutionary precursor to human language affixation. In that study, cotton-top tamarin monkeys were shown to be able to discriminate between artificial words that start with a ‘prefix’ syllable and those that end with the same syllable as a ‘suffix’. Up to now, there is no evidence from any other animal species suggesting an ability to learn patterns similar to the (surface) transformations of human affixation. In **chapter 3**, we trained and tested zebra finches to discriminate sequences constructed to resemble prefixation and suffixation patterns to further investigate this learning ability in nonhuman animals,

‘Algebraic’ rule learning

Studies exploring generalizations of more sophisticated rules, such as ‘algebraic’ rules and the learning of recursive structures are at the core of debates about the uniqueness of the faculty of language. An algebraic rule is defined as an “open-ended” abstract relationship for which we can substitute arbitrary items (Marcus, Vijayan et al., 1999). In a study of human infants, Marcus, Vijayan et al. (1999) exposed 7-months old infants to *XYX* or *XXY* speech stimuli (*X* and *Y* being consonant-vowel syllables) and showed that infants could abstract the underlying regularities and apply them to novel stimuli consisting of new syllables. This ‘algebraic’ generalization by the young infants inspired researchers to explore the presence of such an ability in nonhuman animals. By using similar designs, tests have been performed with monkeys, rats and birds (Hauser & Glynn, 2009; Murphy, Mondragon et al., 2008; Seki, Suzuki et al., 2013; Toro & Trobalon, 2005; van Heijningen, Chen et al., 2013). These animals were found to be able to discriminate different artificial structures. However, the question of whether or not these animals generalized the structures without relying on the physical similarities between stimuli still remains disputable (ten Cate & Okanoya, 2012). Therefore, **chapter 4** deals with this the question by addressing whether zebra finches show the ability of learning and generalizing ‘algebraic’ rules similar to those present in human infants.

Nonadjacent dependencies

Learning natural human languages is not just about generalizing simple and linear structures. A simple sentence like ‘A fish swims’ is not as simple as three words being put together linearly. In this sentence, a morpheme ‘s’ was added spanning the intervening materials (‘fish’ and ‘swim’) to agree with the single number ‘A’. Linguists have suggested that such nonadjacent dependencies might ‘define’ human language when compared to animal vocal communication (Everaert & Huybregts, 2013). Although nonadjacent dependencies are not easy to acquire, even for humans, there is some evidence suggesting that non-human primates are aware of nonadjacent dependencies between sounds (Newport, Hauser et al., 2004; Ravignani, Sonnweber et al., 2013). But experiments involving nonadjacent dependencies that have been done so far were all in primates and the dependencies between elements in these studies all occurred at the edges. In **chapter 5**, we trained and tested zebra finches with an

artificial language consisting of different nonadjacent grammars to address the question whether or not they can learn nonadjacent dependencies and can even learn such dependencies when dependent items occur at arbitrary positions of sequences.

Finally, a summary of the results is given in the last chapter (**6: summary and conclusion**), which discusses the most important findings of this thesis.

References

- Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, 49(6), 1053-1076.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324.
- Berwick, R. C., Okanoya, K., Beckers, G. J., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in cognitive sciences*, 15(3), 113-121. doi: 10.1016/j.tics.2011.01.002
- Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, 11(11), 747-759. doi: 10.1038/Nrn2931
- Bolhuis, J. J., & Wynne, C. D. (2009). Can evolution explain how minds work? *Nature*, 458(7240), 832-833.
- Boughman, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1392), 227-233.
- Comins, J. A., & Gentner, T. Q. (2010). Working memory for patterned sequences of auditory objects in a songbird. *Cognition*, 117(1), 38-53. doi: 10.1016/j.cognition.2010.06.009
- Correia, S. P., Dickinson, A., & Clayton, N. S. (2007). Western scrub-jays anticipate future needs independently of their current motivational state. *Current Biology*, 17(10), 856-861.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. doi: 10.1146/annurev.neuro.22.1.567
- Endress, A. D., Cahill, D., Block, S., Watumull, J., & Hauser, M. D. (2009). Evidence of an evolutionary precursor to human language affixation in a non-human primate. *Biology Letters*, 5(6), 749-751. doi: 10.1098/rsbl.2009.0445
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal cognition*, 13(3), 483-495. doi: 10.1007/s10071-009-0299-8
- Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, 105(3), 577-614. doi: 10.1016/j.cognition.2006.12.014
- Endress, A. D., Nespore, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in cognitive sciences*, 13(8), 348-353. doi: 10.1016/j.tics.2009.05.005
- Everaert, M., & Huybregts, R. (2013). The design principles of natural language. In J. J. Bolhuis & M. Everaerts (Eds.), *Birdsong, speech, and language* (pp. 243-260). Cambridge: The MIT press.
- Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics*, 25(4), 166-177.
- Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in cognitive sciences*, 4(7), 258-267.
- 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(5598), 1569-1579. doi: 10.1126/science.298.5598.1569
- Hauser, M. D., & Glynn, D. (2009). Can free-ranging rhesus monkeys (*Macaca mulatta*) extract artificially created rules comprised of natural vocalizations? *Journal of Comparative Psychology*, 123(2), 161-167. doi: 10.1037/a0015584
- 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(3), B53-64.
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., . . . Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in psychology*, 5.
- Henson, R. N. (1998). Short-term memory for serial order: the Start-End Model. *Cognitive Psychology*, 36(2), 73-137. doi: 10.1006/cogp.1998.0685
- Hitch, G. J., Fastame, M. C., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory*, 13, 247-258.
- Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition*, 97(2), 211-225.
- Janik, V. M., & Slater, P. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59-99.
- Janik, V. M., & Slater, P. J. (1998). Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. *Animal Behaviour*, 56(4), 829-838.
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., . . . Shimizu, T. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6(2), 151-159.
- Kelly, M. H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua: International Review of General Linguistics*, 92, 105-140.

- Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biology Letters*, 6(2), 156-159.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77-80.
- Marslen-Wilson, W. D., & Tyler, L. K. (1998). Trends in cognitive sciences. *Trends in cognitive sciences*, 2(11), 428-435.
- McKenzie, T., Cherman, T., Bird, L. R., Naqshbandi, M., & Roberts, W. A. (2004). Can squirrel monkeys (*Saimiri sciureus*) plan for the future? Studies of temporal myopia in food choice. *Animal Learning & Behavior*, 32(4), 377-390.
- Mulcahy, N. J., & Call, J. (2006). Apes save tools for future use. *Science*, 312(5776), 1038-1040.
- Murphy, R. A., Mondragon, E., & Murphy, V. A. (2008). Rule learning by rats. *Science*, 319(5871), 1849-1851. doi: 10.1126/science.1151564
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, 49(2), 85-117. doi: 10.1016/j.cogpsych.2003.12.002
- 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. *Animal cognition*, 15(2), 155-161. doi: 10.1007/s10071-011-0441-2
- Ohms, V. R., Gill, A., Van Heijningen, C. A., Beckers, G. J., & ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 277(1684), 1003-1009. doi: 10.1098/rspb.2009.1788
- Orlov, T., Amit, D. J., Yakovlev, V., Zohary, E., & Hochstein, S. (2006). Memory of ordinal number categories in macaque monkeys. *Journal of Cognitive Neuroscience*, 18(3), 399-417.
- Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, 404(6773), 77-80. doi: 10.1038/35003571
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1993). Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology*, 26(7), 389-406.
- Pinker, S. (1991). Rules of Language. *Science*, 253(5019), 530-535. doi: 10.1126/science.1857983
- Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Planning for the future by western scrub-jays. *Nature*, 445(7130), 919-921.
- Ravignani, A., Sonnweber, R.-S., Stobbe, N., & Fitch, W. T. (2013). Action at a distance: dependency sensitivity in a New World primate. *Biology Letters*, 9(6), 20130852.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.
- Saffran, J. R., Hauser, M., Seibel, R., Kapfhammer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, 107(2), 479-500.
- 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(1), 27-52.
- Seki, Y., Suzuki, K., Osawa, A. M., & Okanoya, K. (2013). Songbirds and humans apply different strategies in a sound sequence discrimination task. *Frontiers in psychology*, 4, 447. doi: 10.3389/fpsyg.2013.00447
- Spierings, M. J., & ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human.
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1984-1994. doi: 10.1098/rstb.2012.0055
- Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, 14(1), 66-73.
- Toro, J. M., & Trobalon, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & psychophysics*, 67(5), 867-875.
- 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(2), 165-175. doi: 10.1007/s10071-012-0559-x
- Whiten, A., & Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. In P. J. B. Slater, J. S. Rosenblatt, C. Beer, & M. Milinski (Eds.), *Advances in the study of behavior* (pp. 239-283). San Diego: Academic Press.
- Wilbrecht, L., & Nottebohm, F. (2003). Vocal learning in birds and humans. *Mental retardation and developmental disabilities research reviews*, 9(3), 135-148.

Chapter 2^{*}

Zebra finches can use positional and transitional cues to distinguish vocal element strings

^{*} This chapter is based on: Jiani Chen & Carel ten Cate (in press). Zebra finches can use positional and transitional cues to distinguish vocal element strings. *Behavioural processes*

Abstract

Learning sequences is of great importance to humans and non-human animals. Many motor and mental actions, such as singing in birds and speech processing in humans, rely on sequential learning. At least two mechanisms are considered to be involved in such learning. The chaining theory proposes that learning of sequences relies on memorizing the transitions between adjacent items, while the positional theory suggests that learners encode the items according to their ordinal position in the sequence. Positional learning is assumed to dominate sequential learning. However, human infants exposed to a string of speech sounds can learn transitional (chaining) cues. So far, it is not clear whether birds, an increasingly important model for examining vocal processing, can do this. In this study we use a Go-Nogo design to examine whether zebra finches can use transitional cues to distinguish artificially constructed strings of song elements. Zebra finches were trained with sequences differing in transitional and positional information and next tested with novel strings sharing positional and transitional similarities with the training strings. The results show that they can attend to both transitional and positional cues and that their sequential coding strategies can be biased towards transitional cues depending on the learning context.

Introduction

Sequential learning, i.e. learning about the order in which events occur, is of prominent importance in the lives of humans. It enables activities ranging from the production of action sequences to the processing of language. Different theories have been proposed with respect to the cognitive mechanisms involved in sequence learning. One of these, the ‘chaining’ theory, assumes that a sequence of acts or items is remembered by pair-wise association of adjacent elements of a sequence (Henson, 1998): the sequence ABCDE is remembered by A triggering the representation of B, which in turn triggers C, etc.. An alternative theory is that elements are encoded by their ordinal position in relation to the beginning or the end of the sequences: ABCDE is remembered by linking element A to position 1, element B to position 2 etc. (Henson, 1998). Over the years, most evidence is in favor of the positional theory. For example, most chaining models encounter difficulties in dealing with repeated items where two or more responses share the same cue (Henson, 1998, 2001). They also have a problem in handling erroneous responses, which leads to a cascade of subsequent errors (Henson, 1998, 2001). In contrast, the positional theory does not have problems with repeated items nor with erroneous responses. Serial recall experiments in humans have provided support for positional coding (Chiara Fastame, Flude et al., 2003; Endress, Nespor et al., 2009; Henson, 2001), e.g. the phenomenon of serial order intrusion errors, in which an item of one string gets inserted in another one. Such insertions tend to occur at the same ordinal position in the new string as the position they occupied in the original string, suggesting the coding of the item’s position (Chiara Fastame, Flude et al., 2003).

Chaining and positional theories also have a long history in studies of sequence learning in non-human animals (Comins & Gentner, 2010). Animals can also memorize a sequence of events and reproduce ordered lists of artificial items (Orlov, Amit et al., 2006; Orlov, Yakovlev et al., 2000; Terrace, 1987, 1993; Terrace, Son et al., 2003). Again, the positional theory has gained considerable support from studies on species ranging from apes to birds (Comins & Gentner, 2010; Conway & Christiansen, 2001; Endress, Carden et al., 2010; Orlov, Amit et al., 2006; Orlov, Yakovlev et al., 2000; Scarf & Colombo, 2010; Terrace, Son et al., 2003). For instance, starlings could learn to differentiate a string with vocal elements in fixed positions from randomly ordered strings. However, when absolute position cues were removed but sequential information left intact, recognition failed (Comins & Gentner, 2010). This finding is of relevance in the context of comparative research on a particular type of sequence learning, in which vocal elements follow each other with different transitional probabilities. When 8-month-old human infants are exposed to a string of speech syllables consisting of combinations of frequently co-occurring elements, for which there is no cue to their relative position other than the transitional probability they use this information to identify trisyllabic nonsense ‘words’ from such strings (Saffran, Aslin et al., 1996). Hence, in this context young infants are capable of statistical learning, a chaining-like mechanism based on transitional probabilities, that may aid them in natural language learning (Conway & Christiansen, 2001). Cotton-top tamarins, tested in a similar design as human infants, also were sensitive to the transitional information in strings, suggesting that this mechanism may be domain-general and present in animals (Hauser, Newport et al., 2001, see also Kelly & Martin, 1994). This makes it of particular interest to

know whether songbirds, which are excellent model species for comparative studies on vocal processing (Bolhuis & Everaert, 2013; C. ten Cate, 2014), are also able to distinguish strings based on transitional cues of the constituent items.

In this study, we trained zebra finches to discriminate between strings with identical numbers and types of items, but in which the position of, as well as the transitional probabilities between the items differ. Zebra finches are extensively studied as a comparative model for linguistic processing, ranging from studies on the molecular and neural mechanisms of vocal learning (Bolhuis, Okanoya et al., 2010), up to the processing of speech sounds (Ohms, Gill et al., 2010) and artificial grammar learning (van Heijningen, Chen et al., 2013; van Heijningen, de Visser et al., 2009). That zebra finches can learn sequences of song elements is well established in the context of song learning, where young birds copy songs from song tutors. They may do so by learning chunks of elements (C. ten Cate & Slater, 1991; Williams & Staples, 1992), but this observation provides no insight in the mechanism involved. A recent study (Lipkind, Marcus et al., 2013) showed that zebra finches might copy element sequences by first concentrating on bigrams, which may indicate the use of a chaining-like mechanism during song learning. However, adult zebra finches can also learn to discriminate strings of elements in which an odd element is embedded either early or late in a series of identical elements (Verzijden, Etman et al., 2007) – a task that can be solved by positional learning, but not by chaining. So, such circumstantial evidence suggests that zebra finches can use both positional and chaining mechanisms to memorize element sequences, but thus far no study tested this explicitly.

In the current experiment, zebra finches could use both transitional and positional information to discriminate two strings consisting of song elements arranged in different patterns. Given that positional learning is the most likely strategy for sequential learning, we designed our experiments in such a way that solving the task by using transitional information would be encouraged. After a discrimination training on these strings, we tested the bird's sequential coding strategies by giving probe tests with novel sequences that contained different degrees of chaining and positional information. The results show that zebra finches used both types of information.

Material and method

Subject and apparatus

Eight zebra finches (4 males and 4 females, age: $226d \pm 15$ SEM) were individually housed in operant conditioning cages (70×30×45 cm) in sound attenuated chambers. The cages were made of wire mesh with a plywood back wall that contained a food hatch and two red pecking sensors. Each sensor contained a red LED that indicated the activated stage of the sensor. Five perches were mounted between the back and front side of the cage to enable hopping behavior. The sensors and the food hatch could be reached from the middle perch. The birds remained in the operant conditioning cages until they completed the training and tests. They were kept on a 10.5 – 13.5 dark – light schedule (similar to the dark – light condition of the breeding colony in which they were housed previously) Drinking water and cuttlebone were available *ad libitum* during the whole period of the experiments. The food intake was measured

daily to keep track of a sufficient food intake. The study was conducted according to Association for the Study of Animal Behavior guidelines on animal experimentation as well as to the Dutch law on animal experimentation, and approved by the Leiden University Committee for animal experimentation (DEC).

Acoustic stimuli were delivered from a loudspeaker located above the cage and were calibrated to a pressure level of 70 dB. A custom-designed controller controlled the pecking sensors, food hatch, chamber light, auditory stimuli presentation and also registered the responses of a subject.

Stimuli

Stimuli were constructed from ramped and RMS equalized zebra finch song elements. Song elements were arranged in a linear way with 40ms pause inserting in between two adjacent elements, which is comparable to the duration of pauses between elements in natural zebra finch songs. The duration of all stimuli was less than 1.6s (see Fig. 1 for an example). Natural zebra finch songs consist of a series of song elements that may differ in type or order among different males (Zann & Bamford, 1996). There is no indication that single elements or element combinations carry specific meanings and hence these elements are suitable for constructing artificial strings. We selected six element types, based on optimal discriminability, from our zebra finch song database. To eliminate pseudo effects due to any arbitrary parameter of the song elements, the assignments of the element types were shuffled for each subject. For instance, a string ‘abcdef’ has different combinations of element types for different individuals.

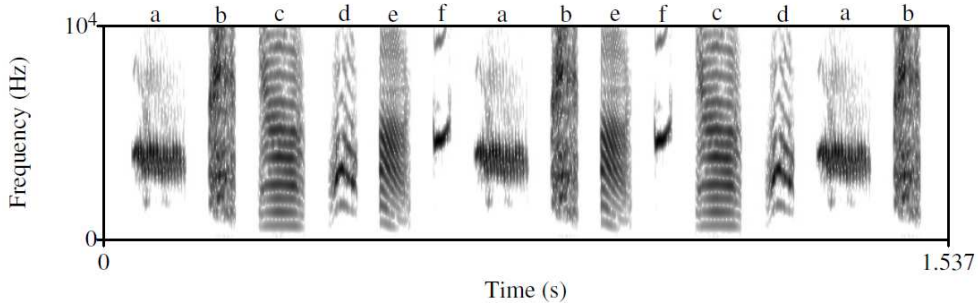


Fig. 1 Spectrogram of a Go stimulus for one bird.

Song elements were assigned to different ordinal positions to construct stimuli. Besides differences between the stimuli in the position of elements, the chaining regularities also differed between different sets of sequences. ‘Set A’ stimuli always contained ‘cd’ and ‘ef’ bigrams while ‘Set B’ stimuli always contained ‘df’ and ‘ce’ bigrams (Table 1 and 3). In order to encourage the birds to take the full string structure into account when memorizing the strings, all training sequences started and ended with an ‘ab’ bigram. This made it more difficult to differentiate between the strings by just attending to the very first or the final elements of the strings.

Procedure

To familiarize the birds with the Go/Nogo task (appendix Fig. 1), they were first trained to discriminate a natural song (Go stimulus) from a pure tone (Nogo stimulus). Pecking on the first sensor initiated a playback of either a Go or a Nogo stimulus. If the bird pecked the second sensor after hearing a Go stimulus it got 10s access to food. Conversely, if the bird responded to the Nogo stimulus, it would receive a mild punishment by switching off the light of the cage for 15s. If the bird did not response within 6s after it pecked the first sensor, no reinforcement would be provided. This phase lasted until the birds' responses reached the criterion of responding to >75% of the Go stimuli and to <25 % of the Nogo stimuli, calculated on the performance over a whole day, for at least two consecutive days. Hereafter the training was switched to the artificial stimuli. For this training, the learning criterion was set at responding > 75% to each Go stimulus and < 25 % to each Nogo stimulus for three consecutive days. After reaching this criterion, a test phase started. Training stimuli were presented and reinforced on 80% of the trials to prevent extinction of the pecking behavior. Test stimuli were presented on 20% of the trials and were never reinforced. They were presented in a random order. When each test stimulus had been presented in 40 trials, tests were stopped and the birds were switched to the next phase. These 40 trials of each stimulus were used in the statistical analysis.

Table 1 Training and test stimuli in Experiment 1

Experiment 1	Set A		Set B	
Training	Go	abcdefabefcdab	Nogo	abdfcedfabceab
Test	A1	ab <u>efcd</u> ab <u>cd</u> efab	B1	ab <u>cd</u> ef <u>ce</u> ab <u>bd</u> fab
	A2	ab <u>efcd</u> efab <u>bc</u> dab	B2	ab <u>cd</u> efab <u>bd</u> f <u>ce</u> ab

Letters indicate different song elements. Underlined: elements in identical positions to the Go-string; bold: elements in identical position to the Nogo string.

Experiment 1

In Experiment 1, the birds were first trained to discriminate two different stimuli, one from Set A and one from Set B, which were presented as Go and Nogo stimulus respectively (Table 1). The Go and the Nogo stimuli differed in the ordinal positions of elements (except for the ‘ab’ elements at the edges), as well as in the transitional probabilities of the elements (see Table 2). After reaching the training criterion, four non-reinforced test stimuli were mixed between the training stimuli (A1, A2, B1 and B2, Table 1). The test stimuli were identical to the training stimuli of their sets with respect to the transitional probabilities (Table 2). However, the positional information differed substantially from the training stimuli. Apart from the identical ‘ab’ bigrams at the beginning and the end, the test stimuli A1 and B1 both had two elements in positions identical to the Go stimulus and two elements in positions identical to the Nogo stimulus (Table 1). For test stimuli A2 and B2, transitional and positional information were in conflict: they contained the transitional information of their own set, but the ordinal position of elements was more like the training stimulus of the other set. For example, test stimulus A2

shared some unique bigrams with the Go training stimulus, such as ‘cd’, ‘ef’ etc., while there were no unique bigrams shared between A2 and the Nogo training stimulus (Table 2). With respect to the positional information, it had six elements at the positions that were identical to the Nogo training sequence, whereas only two elements were at positions identical to the Go training sequence (excluding the non-informative ‘ab’-bigrams at both edges - Table 1). If the training sequences are distinguished using positional information only, we expect that the birds will not differentiate in their responses to stimuli A1 and B1, while they should treat stimulus A2 more like a Nogo stimulus and B2 more like a Go-stimulus. If the birds use transitional information, they should treat A1 and A2 like Go-stimuli and B1 and B2 like Nogo-stimuli.

Table 2 Matrix of bigrams occurred in the training and test stimuli of Experiment 1

		Following					
		a	b	c	d	e	f
Preceding	a	0/0	3/3	0/0	0/0	0/0	0/0
	b	0/0	0/0	1/1	0/1	1/0	0/0
	c	0/0	0/0	0/0	2/0	0/2	0/0
	d	1/0	0/0	0/0	0/0	1/0	0/2
	e	0/1	0/0	0/0	0/1	0/0	2/0
	f	1/1	0/0	1/1	0/0	0/0	0/0

Transition matrix of element bigrams for the training and test stimuli. The numbers before and after the slash indicate numbers of the bigrams present in the Set A and Set B sequences respectively. Strings within Set A and Set B were identical in transitional probabilities between elements.

Experiment 2

All birds from Experiment 1 continued with Experiment 2 after finishing the tests of Experiment 1, using the same training criterion and test procedures as for Experiment 1. In Experiment 2, the training and test stimuli from Experiment 1 were combined and together used as training stimuli (Table 3). The stimuli from Set A served as Go stimuli and those from Set B as Nogo stimuli.

Stimuli in Test 1 of this experiment were identical to the training stimuli with respect to chaining regularities, apart from one ‘fc’ transition becoming an ‘fa’ transition in both test stimuli (Test 1, Table 3). Apart from the final ‘ab’ bigram, six elements of stimulus A3 can be found in positions present in one or two of the Go stimuli and five elements in positions present in one or two of the Nogo stimuli. For stimulus B3 this is the other way round. The rationale for this test was to examine whether the extended training set, in which transitional information remained identical for the stimuli of one set, but the positional information varied may have a positive effect on the use of transitional information. If so, the birds should show clear discrimination between the two test stimuli. If, on the other hand, the birds had learned all individual training stimuli by remembering their element positions, they should show no or very little differentiation between the test strings.

Test 2 essentially addressed the same question, but provided more systematic variation in element position (Test 2, Table 3). The stimuli in this test are identical to the first Go (Go1) and Nogo (Nogo1) stimuli in the training with respect to the exact sequence of elements, but with a systematic movement of the elements by two positions. For example, the first two

elements of the Go1 stimulus are moved to the end of the sequence for the first test stimulus (A4), and next the first two elements of A4 were moved to the end to make the second test stimulus (A5) and so on. In this way, some chaining cues that occur in the training stimuli were broken. Chaining items like ‘ba’ never occurred in neither Go nor Nogo training stimuli. Also, several of these stimuli did not have an ‘ab’ sequence at the beginning or end, which might disrupt remembering the positions of items by using the begin or end of a string as anchor positions.

Table 3 Training and test stimuli in Experiment 2

Experiment 2	Set A		Set B	
Training	Go1	abcdefabefcdab	Nogo1	abdfcdfabceab
	Go2	abefcdabcedfab	Nogo2	abcedfcebdfab
	Go3	abefcdefabacdab	Nogo3	abcedfabdfceab
Test 1	A3	cdefabacdabefab	B3	ceabdfabcedfab
Test 2	A4	cdefabefcdabab	B4	dfcedfabceabab
	A5	efabefcdababcd	B5	cedfabceababdf
	A6	abefcdababcdef	B6	dfabceababdfce
	A7	efcdababcedfab	B7	abceababdfcedf
	A8	cdababcedfabef	B8	ceababdfcedfab
	A9	ababcedfabefcd	B9	ababdfcedfabce
Test 3	A10	cdef	B10	cedf
	A11	efcd	B11	dfce

If the birds used the transitional regularities to distinguish the sets, they might also be able to discriminate between shorter segments differing in transitional information. Stimuli of Test 3 consisted of four items only, but the segments kept the transitional regularities of either the Go or the Nogo training stimuli. They contained the most diagnostic bigrams for discriminating the stimuli by differences in element transitions.

Results

Experiment 1

All birds reached the training criterion after, on average, 6343 training trials (± 878 SEM). The proportion of Go-responses to the four test stimuli (A1, A2, B1 and B2, 40 trials per stimulus, results in Table 4) showed no significant difference at the group level ($n = 8$, $p = 0.125$, $\chi^2(3) = 5.734$, Friedman test). At the individual level, five out of eight birds showed more Go-responses to the A1 stimulus than to B1, of which four were significantly different in the number of responses to the two stimuli (chi-square test with Holm-Bonferroni correction for eight comparisons, Table 4). The significant higher proportion of responses to the A1 sequence indicate that the birds may use chaining cues for making the discrimination (Table 4). One of the birds (bird 8) responded significantly less to the A1 stimulus than to the B1 stimulus.

Table 4 Proportion of Go-responses of individual birds in Experiment 1

Bird	Training		Test							
	Go	Nogo	A1	B1	<i>p</i>	Transition	A2	B2	<i>p</i>	Transition
1	0.98	0.04	0.98	0.33	<0.001*	+	0.90	0.30	<0.001*	+
2	0.81	0.22	0.80	0.23	<0.001*	+	0.70	0.43	0.024	
3	0.91	0.10	0.78	0.45	0.006*	+	0.70	0.50	0.110	
4	0.93	0.11	0.88	0.73	0.162		1.00	0.53	<0.001*	+
5	0.88	0.13	0.90	0.63	0.008*	+	0.55	0.58	1.000	
6	0.79	0.06	0.43	0.55	0.371		0.60	0.33	0.024	
7	0.77	0.05	0.48	0.73	0.039		0.58	0.43	0.264	
8	0.88	0.07	0.35	0.88	<0.001*	-	0.35	0.80	<0.001*	-
Mean	0.87	0.10	0.70	0.57			0.67	0.49		
SEM	0.03	0.02	0.09	0.08			0.07	0.06		

‘+’ symbol indicates a possibility of having used chaining information. ‘-’ symbol indicates little possibility of using transitional information. ‘*’ symbol indicates significant difference after Holm-Bonferroni correction.

Six birds responded to A2 stimulus with a higher ratio than to the B2 sequence and two of them reached a significant level (chi-square test with Holm-Bonferroni correction for eight comparisons, Table 4). One bird (again bird 8) had a significantly higher responding ratio to B2 than to A2.

Experiment 2

Test 1

Table 5 Proportion of Go-responses of individual birds in Experiment 2 – Test 1

Bird	Training		Test		
	Go 1 – 3 (Mean ± SEM)	Nogo 1 – 3 (Mean ± SEM)	A3	B3	<i>p</i>
1	0.93 ± 0.003	0.03 ± 0.007	0.58	0.00	<0.001*
2	0.84 ± 0.018	0.11 ± 0.020	0.66	0.17	<0.001*
3	0.92 ± 0.009	0.08 ± 0.019	0.54	0.07	<0.001*
4	0.96 ± 0.003	0.02 ± 0.006	0.40	0.05	<0.001*
5	0.92 ± 0.012	0.23 ± 0.006	0.85	0.24	<0.001*
6	0.88 ± 0.009	0.02 ± 0.003	0.78	0.05	<0.001*
7	0.91 ± 0.006	0.01 ± 0.003	0.58	0.03	<0.001*
8	0.78 ± 0.034	0.17 ± 0.019	0.78	0.53	0.035*

‘*’ symbol indicates significant difference after Holm-Bonferroni correction.

All birds reached training criterion after, on average, 5412 training trials (± 1866 SEM) and results of the test showed that all of them discriminated between the two test stimuli (Table 5). The proportions of Go-responses of these birds to stimulus A3 were significantly higher than responses to stimulus B3 (chi-square test with Holm-Bonferroni correction for 8 comparisons, Table 5).

Test 2

In Test 2, the birds’ proportion of Go-responses to the test stimuli from Set A (calculated as average value per bird over all stimuli), was significantly higher than the proportion of Go-responses to Set B (A4 – A9: 0.68 ± 0.05 SEM, B4 – B9: 0.15 ± 0.06 SEM, $z = -2.52$, $p = 0.012$, $n = 8$, Wilcoxon signed-rank test). The responses to the different test stimuli within a set varied significantly for Set A ($p = 0.001$, $\chi^2(5) = 21.097$, Friedman test) but not for Set B ($p = 0.09$, $\chi^2(5) = 9.523$, Fig. 2). The response proportion to stimulus A6 was the highest

and responses to stimuli A4, A5 and A9 were similar while responses to A7 and A8 were relatively lower. Although responses to different stimuli in Set B were slightly different, the proportions of responses were all low.

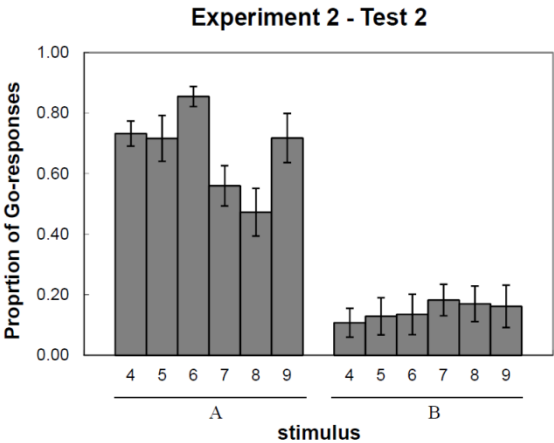


Fig. 2. Mean ($\pm$ SEM, $n = 8$) proportion of Go- responses to the test stimuli in Test 2. Numbers under the horizontal abscissa indicate the individual stimuli in the Set A and Set B. Responses between Set A and Set B were significantly different.

Test 3

In this test, birds discriminated among the four-item segments according to their transitional regularities (Fig. 3). The average proportion of Go-response to the two segments with ‘cd’ and ‘ef’ transitions was significantly higher than the average response to the segments with ‘ce’ and ‘df’ transitions (Set A vs Set B: $z = -2.10$, $p = 0.035$, $n = 8$, Wilcoxon signed-rank test).

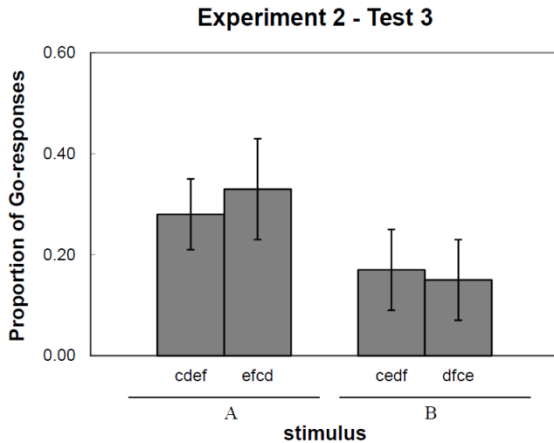


Fig. 3. Mean ($\pm$ SEM, $n = 8$) proportion of Go-responses to the test stimuli in Test 3. Responses between Set A and Set B were significantly different.

Discussion

In Experiment 1, zebra finches that had a higher response to test stimulus A1 than to B1 (with A1 identical in transitional information to the Go and B1 to the Nogo stimulus, but with both being equally similar to both Go and Nogo stimuli with respect to positional information) must have used transitional information to differentiate the test stimuli. When the test stimuli shared more positional similarity with one training stimulus and more transitional similarity with the second training stimulus (A2 and B2), fewer birds showed significant discrimination, which suggests that most birds combined transitional and positional information for their discrimination. Nevertheless, overall the transitional information seems to dominate string memorization, both on average as well as for most birds individually. However, the finding that one bird (bird 8) responded more to stimuli B1 and B2 than to stimuli A1 and A2 suggests that it used positional information, most likely the presence or absence of the ‘c’ element in the third position. This element is in the third position of the ‘Go’ stimulus, but for the test sounds it is only in that position for the stimulus set that shared its transitional probabilities with the Nogo training strings.

In Experiment 2, the test stimuli of Experiment 1 were included in the training sets. This should stimulate the birds to give more attention to transitional cues (identical among the stimuli of a set) than to positional ones (different among the stimuli of a set). The results of both test 1 and 2 of Experiment 2 show that all birds now discriminated the test stimuli according to their transitional information. However, as the discrimination between the stimuli in test 1 is reduced compared to the training, it cannot be concluded that the birds responded exclusively to transitional probabilities and completely ignored positional cues. This is also supported by the outcome of the second test of Experiment 2. Again, the results show that birds discriminated between stimuli from the two sets, with a higher proportion of Go-responses to the stimuli sharing most transitional similarities with the Go training stimuli. However, the responses to individual stimuli within Set A varied from one to another, which might be due to the additional use of positional cues. For example, the zebra finches responded strongest to test stimulus A6, which shared the position of its initial elements with training stimuli Go 2 and 3.

In the third test of Experiment 2, the zebra finches were able to discriminate segments of the training strings according to their chaining regularities, although the level of responding to the test strings originating from the Go-stimulus was reduced compared with responding to the Go-stimuli. Nevertheless, the birds recognized the difference between the strings, even though the length of segments was much shorter than the training strings, and lacked part of the transitions differentiating these strings (e.g. ‘bd’ and ‘be’). Zebra finches were thus able to categorize the different stimuli by using only part of the acoustic features of the element strings.

Our results differ from those of previous experiments in which animals were trained with acoustic strings of species-specific sounds containing both positional and transitional information (Comins & Gentner, 2010; Endress, Carden et al., 2010). These showed that both starlings and chimpanzees had a learning bias to use positional information in the encoding of sequences, with the starlings having great difficulties in using transitional information to distinguish among strings (Comins and Gentner 2010). This contrast with our finding that zebra

finches use transitional information, might be due to a different training schedule, as we deliberately modelled the training stimuli in such a way that it might encourage the use of transitions rather than, or in addition to, ordinal positions to memorize the training stimuli. However, our results may also reflect a species difference, with zebra finches being better able to distinguish sequences based on transitional regularities than the other species, although it is hypothesized that starlings might also have both mechanisms available (Comins and Gentner 2010). The sensitivity to transitions in the zebra finch fits with recent evidence from song learning studies showing that zebra finch song learning may be guided by first learning about element transitions (Lipkind, Marcus et al., 2013).

In general, sequential learning can make use of both positional and transitional or chaining information, with the use of positional information being most prominent and most universal and limited evidence for using transitional information (Comins & Gentner, 2010; Conway & Christiansen, 2001; Endress, Carden et al., 2010; Hauser, Newport et al., 2001; Orlov, Amit et al., 2006; Orlov, Yakovlev et al., 2000; Scarf & Colombo, 2010; Terrace, Son et al., 2003; Toro & Trobalon, 2005). Our findings suggest not only that the zebra finches can use transitional information together with positional information, but also that they are able to bias their sequential coding strategy depending on what might be most profitable given the context.

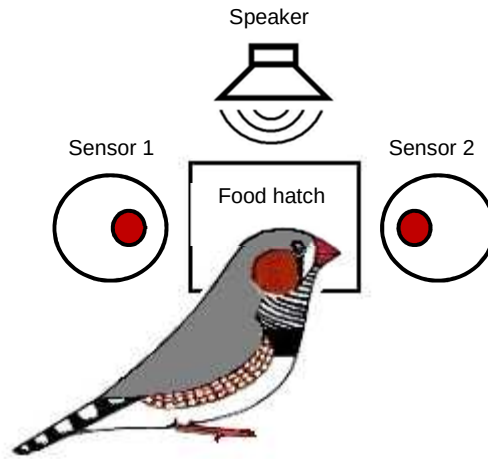
The clear presence of using transitional information in the zebra finches is of particular interest from a comparative perspective. A fundamental process during language acquisition is to segment words from fluent speech. One way in which humans accomplish this task is by attending to transitional cues in continuous speech (Aslin, Saffran et al., 1998; Saffran, Aslin et al., 1996). Although birds do not have language and are phylogenetically distant from humans, they also seem able to do this. This suggests that the ability to use transitional information to distinguish vocal strings is a more general cognitive capacity that may have been recruited in the evolution of language.

References

- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324.
- Bolhuis, J. J., & Everaert, M. (2013). *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain*. Cambridge: MIT Press.
- Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, 11(11), 747-759. doi: 10.1038/Nm2931
- Chiara Fastame, M., Flude, B., & Hitch, G. J. (2003). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory*, 13(3-4), 247-258.
- Comins, J. A., & Gentner, T. Q. (2010). Working memory for patterned sequences of auditory objects in a songbird. *Cognition*, 117(1), 38-53. doi: 10.1016/j.cognition.2010.06.009
- Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in cognitive sciences*, 5(12), 539-546.
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal cognition*, 13(3), 483-495. doi: 10.1007/s10071-009-0299-8
- Endress, A. D., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in cognitive sciences*, 13(8), 348-353. doi: 10.1016/j.tics.2009.05.005
- 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(3), B53-64.
- Henson, R. N. (1998). Short-term memory for serial order: the Start-End Model. *Cognitive Psychology*, 36(2), 73-137. doi: 10.1006/cogp.1998.0685
- Henson, R. N. (2001). Serial order in short-term memory. *The Psychologist*, 14(2), 70-73.
- Kelly, M. H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua: International Review of General Linguistics*, 92, 105-140.
- Lipkind, D., Marcus, G. F., Bemis, D. K., Sasahara, K., Jacoby, N., Takahasi, M., . . . & Tchernichovski, O. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*. 498(7452), 104-108
- Ohms, V. R., Gill, A., Van Heijningen, C. A., Beckers, G. J., & ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 277(1684), 1003-1009. doi: 10.1098/rspb.2009.1788
- Orlov, T., Amit, D. J., Yakovlev, V., Zohary, E., & Hochstein, S. (2006). Memory of ordinal number categories in macaque monkeys. *Journal of Cognitive Neuroscience*, 18(3), 399-417.
- Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, 404(6773), 77-80. doi: 10.1038/35003571
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.
- Scarf, D., & Colombo, M. (2010). A positional coding mechanism in pigeons after learning multiple three-item lists. *Animal cognition*, 13(4), 653-661.
- ten Cate, C. (2014). On the phonetic and syntactic processing abilities of birds: from songs to speech and artificial grammars. *Current Opinion in Neurobiology*, 28, 157-164.
- ten Cate, C., & Slater, P. (1991). Song learning in zebra finches: how are elements from two tutors integrated? *Animal Behaviour*, 42(1), 150-152.
- Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature*, 325(6100), 149-151. doi: 10.1038/325149a0
- Terrace, H. S. (1993). The phylogeny and ontogeny of serial memory: list learning by pigeons and monkeys. *Psychological Science*, 4(3), 162-169.
- Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, 14(1), 66-73.
- Toro, J. M., & Trobalon, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & psychophysics*, 67(5), 867-875.
- 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(2), 165-175. doi: 10.1007/s10071-012-0559-x
- 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. *Proceedings of the National Academy of Sciences*, 106(48), 20538-20543. doi: 10.1073/pnas.0908113106

- Verzijden, M. N., Etman, E., van Heijningen, C., van der Linden, M., & ten Cate, C. (2007). Song discrimination learning in zebra finches induces highly divergent responses to novel songs. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 274(1607), 295-301.
- Williams, H., & Staples, K. (1992). Syllable chunking in zebra finch (*Taeniopygia guttata*) song. *Journal of Comparative Psychology*, 106(3), 278.
- Zann, R. A., & Bamford, M. (1996). *The zebra finch: a synthesis of field and laboratory studies* (Vol. 5): Oxford University Press Oxford.

Appendix



Appendix Fig. 1 Diagram of the operant conditioning apparatus. During the learning task, zebra finch learns to peck the first sensor to trigger a playback of an acoustic stimulus. Pecking the second sensor according to the sound playback leads to either a food reward or a mild punishment of couple seconds of darkness.

Chapter 3^{*}

Zebra finches are able to learn prefixation and suffixation patterns

^{*} This chapter is based on: Jiani Chen, Naomi Jansen & Carel ten Cate (submitted).

Abstract

Adding an affix to transform a word is common across the world languages, with the edges of words more likely to carry out such a function. However, detecting affixation patterns is also observed in learning tasks outside the domain of language, suggesting that the underlying mechanism from which affixation patterns have arisen may not be language- or even human-specific. We addressed whether a songbird, the zebra finch, is able to discriminate between, and generalize about, affixation patterns. Zebra finches were trained and tested in a Go/Nogo paradigm to discriminate artificial song element sequences resembling prefixed and suffixed ‘words’. The ‘stems’ of the ‘words’, consisted of different combinations of a triplet of song elements, to which a fourth element was added as either a ‘prefix’ or a ‘suffix’. After training, the birds were tested with novel stems, consisting of either rearranged familiar element types or of novel element types. The birds were able to generalize the affixation patterns to novel stems with both familiar and novel element types. Hence the discrimination resulting from the training was not based on memorization of individual stimuli, but on a shared property among Go or Nogo stimuli, i.e. affixation patterns. Remarkably, birds trained with suffixation as Go pattern showed clear evidence of using both prefix and suffix, while those trained with the prefix as the Go-stimulus used primarily the prefix. This may imply an interesting parallel to the asymmetry in the type of affixation preferred in human languages.

Introduction

Understanding how human language may have arisen is still one of science's greatest challenges. Hypotheses about its evolutionary origin are linked to debates about the properties that make it unique and those that are shared with other animals (Fitch, 2010; Hauser, Chomsky et al., 2002). Comparative studies in non-human animals may thus provide a window on how linguistic complexity may have arisen and what its precursors might have been. Such studies can be directed at related taxa, like apes and monkeys. Alternatively, one can examine the presence of 'linguistic abilities' in more distantly related groups in which relatively complex and structured vocalizations evolved independently, such as songbirds. Songbirds show striking cognitive, neural, and genetic similarities with humans in vocal perception, production and auditory-vocal learning (e.g. Berwick, Okanoya et al., 2011; Bolhuis, Okanoya et al., 2010; Doupe & Kuhl, 1999; Ohms, Gill et al., 2010; ten Cate & Okanoya, 2012). For this reason they are seen as excellent comparative models to explore mechanisms that might have been at the basis of language evolution (Bolhuis & Everaert, 2013). In the current study we also use a songbird species, the zebra finch, and examine whether it is able to discriminate between, as well as to generalize, 'affixation' patterns.

Among the components of language, one of the most notable aspects is morphological transformation. Over decades, studies of inflectional morphology have been at the center of important debates about the fundamental nature of cognitive processes (Endress & Hauser, 2011). Affixation in languages, like the English past tense, provides a crucial case in establishing the view of mental computation as rule-based manipulation of symbol systems (Marslen-Wilson & Tyler, 1998). The great majority of English verbs form their past tense by adding an -ed affix to an unchanged stem. Knowledge of affixation rules plays an important role in learning vocabulary and in language development (Kuczaj, 1977; Mochizuki & Aizawa, 2000; Nagy, Diakidoy et al., 1993). Adding an affix to transform a word occurs quite often across languages. Interestingly, the edges of the words are more likely to carry out the grammatical functions; an affix in the first position (prefix) or in the last one (suffix) is much more frequent than affixes in other positions (Endress & Hauser, 2011; Endress, Nespors et al., 2009). This bias is not only observed in languages. Learning in serial memory tasks also showed that the edge positions of a sequence can be recalled more accurately (Endress, Carden et al., 2010; Henson, 1998, 1999; Hitch, Burgess et al., 1996; Wright, Santiago et al., 1985). Also in artificial language learning, participants were found to generalize regularities reliably at the edges but not in the middle of acoustic sequences (Endress & Mehler, 2010; Endress, Scholl et al., 2005). These examples suggest that prefixation and suffixation patterns are relatively easy to learn and that the edge-based positional learning competence could be an example of what Endress, Nespors, et al., (2009) called "perceptual and memory primitives"; phylogenetically pre-existing cognitive mechanisms which constrain rule-based learning in language acquisition and guide language evolution. If so, it raises the question to which extent the edge-based positional learning competence is shared with non-human animals and whether they can learn affixation rules.

Studies of sequential memorization in general have shown that the edge(s) of a sequence can be recalled better in birds and monkeys (Comins & Gentner, 2010; Endress, Carden et al., 2010; Orlov, Yakovlev et al., 2000; Terrace, Son et al., 2003; Wright, Santiago et al., 1985). In an artificial language learning experiment Endress et al. (2010) showed that chimpanzees also encode the edges of sequences better than the other positions in the sequences, similar to adult humans in the same experiment. Such experiments suggest that animals might also have the ability to recognize and learn affixation patterns, and in a pioneering study (Endress, Cahill et al., 2009), cotton-top tamarin monkeys were exposed to a set of human speech syllables ('stems') that were either preceded or followed by the affix syllable 'shoy'. When subsequently presented with novel stems, the tamarins discriminated between words starting with shoy as a 'prefix' and those that end with the same syllable as a 'suffix', demonstrating that they generalized the underlying affixation rule. Up to now, there is no evidence of such an ability from other animal species. Given the above mentioned similarities between songbirds and humans in vocal processing and also because birds show evidence of at least some, albeit simple, rule learning when trained and tested with strings of elements that are artificially structured according to different algorithms (e.g. Gentner, Fenn et al., 2006; Seki, Suzuki et al., 2013; van Heijningen, Chen et al., 2013; van Heijningen, de Visser et al., 2009), they are promising candidates to examine the presence of affixation learning as a more wide spread cognitive ability that is not specific to language nor to humans.

In the current study, we trained and tested zebra finches in a Go/Nogo paradigm to discriminate sequences constructed to resemble prefixation and suffixation patterns. In doing so, we concentrated on the edges of the sequences, in line with the studies that identified those as being the most salient parts of sequences and also the ones being used most for affixation in languages. We do not want to claim that such an experiment demonstrates the presence of the full formal notion of affixations in a non-human animal. What it can show is whether birds have the competence to detect surface transformations similar to different affixation patterns (prefix and suffix) and link these to different 'meanings' – in this case either a food reward or a mild punishment. The linkage of each affixation pattern to a different reward is also an advantage of using a Go/NoGo paradigm over a habituation paradigm as used in the tamarin study (Endress, Cahill et al., 2009). The habituation paradigm can tell whether animals spontaneously detect a change in a pattern, but detecting such a change is not linked to any consequence. The Go/Nogo not only tests whether the animals detect a difference, but also whether they can link this to a difference in consequences, analogous to human infants that have to learn over time how different affixations alter word meanings.. We also examine whether zebra finches can learn both prefixation and suffixation patterns equally well, or have a bias to be more sensitive to one over the other, as has been suggested for human languages (Cutler, Hawkins et al., 1985; Dryer, 2005; St Clair, Monaghan et al., 2009). We show that the zebra finches are able to learn both regularities. Remarkably, birds that had been trained with prefixation as Go pattern used predominantly the prefix to make their discrimination while birds trained with suffixation as Go pattern used both prefix and suffix.

Material and method

Subject and apparatus

Twelve zebra finches (6 males and 6 females) from the Leiden University breeding colony were trained and tested individually in sound attenuated chambers. None of the birds had previous experience with any kind of experiment. Six birds participated in Experiment 1; all 12 participated in Experiment 2. The experiment was conducted by using the Go/Nogo paradigm in an operant conditioning cage described earlier (van Heijningen, Chen et al., 2013). A fluorescent tube on the top of the box emitted daylight spectrum light on a 13.5 L: 10.5 D schedule. Sounds were played through a loudspeaker, attached above the box, at approximately 70 dB. Subjects gained access to food for 10 seconds after they responded to positive sound playback (Go stimuli). Conversely, if subjects responded to negative sound playback (Nogo stimuli), the light of the chamber was switched off for 15 second to indicate the error.

Stimuli

The ‘words’ used in this study consisted of artificially edited sequences consisting of four song elements. These elements were obtained from natural zebra finch songs (undirected songs) originating from our zebra finch song database. Seven elements, ‘flat,’ ‘stack,’ ‘trill,’ ‘downslide,’ ‘high,’ ‘curve,’ and ‘noisy’ (see Fig 1 for examples), were selected based on optimal discriminability. They were ramped and RMS equalized.

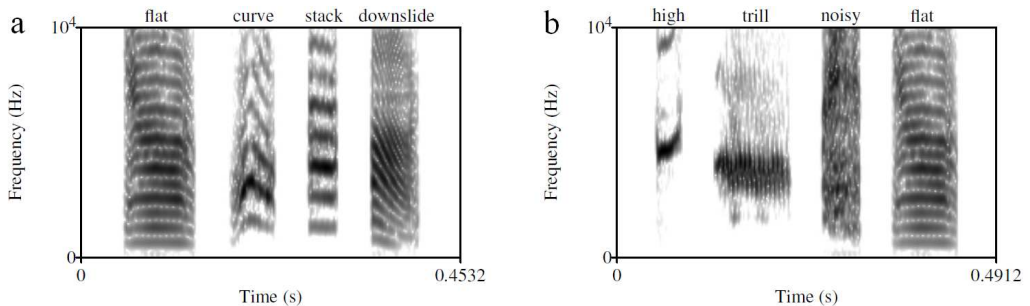


Fig. 1 Spectrograms of stimuli GABC for two different birds (a and b).

Two types of regularities, prefixation and suffixation, were used to construct the stimuli (Table 1). The ‘stems’ of the training stimuli were triplets constructed from three different element types ‘A, B and C’ in different combinations. Each element type occurred in every possible slot over the triplets. A fourth element type ‘G’ was only used as either a ‘prefix’ or a ‘suffix’. To eliminate pseudo effects due to an arbitrary parameter of the sounds, the element assignments were shuffled for the subjects, for instance, element ‘A’ could be ‘curve’ for one bird and ‘trill’ for another bird (Fig. 1). In Experiment 1, birds were first trained with three Go and three Nogo stimuli, each consisting of different combinations of an A, B and C element, and either preceded or followed by the affix (‘G’). The test stimuli of Experiment 1 were constructed by rearranging the element combinations of the stems. In Experiment 2, the training set included

the stimuli for the training as well as those for testing that had been used in Experiment 1. Testing occurred with stimuli in which the stems were formed by the novel element types ‘D, E, and F’, which never occurred in the training phase.

For each stimulus, 40 ms of silence was inserted between consecutive elements and 50 ms of silence was added at the start and the end to avoid acoustic distortions during playback. The training stimuli followed either a prefixation or a suffixation pattern. For half of the birds, the G-prefix predicted Go stimuli and the G-suffix the Nogo stimuli (Table 1, Condition 1), and vice versa for the other half of the birds (Table 1, Condition 2). The test stimuli were constructed by adding the G-suffix or –prefix to novel stems.

Procedure

To familiarize the birds with the Go/Nogo task, they were first trained to discriminate a natural song (Go stimulus) from a pure tone (Nogo stimulus). When their responses reached the training criterion (>75% response to Go stimuli and <25% response to Nogo stimuli) for at least two consecutive days, they were switched to the next phase of training, in which the experimental stimuli were presented.

Experiment 1

This experiment tested whether the birds were able to generalize the affixation patterns of the training stimuli when these were presented in combination with novel stems composed with familiar elements. Six birds (Group 1) were trained with three pairs of stimuli and subsequently tested with another three pairs of stimuli constructed from familiar element types but in novel combinations (Table 1). The tests started after the birds reached the training criterion to every training stimulus for at least three consecutive days.

In the tests, test stimuli were not reinforced to avoid additional learning. Every test contained 40 presentations of each test stimulus. However, to prevent extinction of the pecking behavior, only 20% of all stimulus presentations were test stimuli. The other 80% of stimulus presentations consisted of the reinforced training stimuli.

Table 1 Stimuli of Experiment 1 and 2

Condition	Experiment 1 (Group 1)			Experiment 2 (Group 1 & 2)			
	Training			Training			
1	Go	Nogo	Test	Go	Nogo	Test 1	Test 2
2	Nogo	Go		Nogo	Go		
Stimulus			GACB	GABC	ABCG	GDEF	
			GBAC	GBCA	BCAG	GEFD	
	GABC	ABCG	GCBA	GCAB	CABG	GFDE	ABC
	GBCA	BCAG	ACBG	GACB	ACBG	DEFG	DEF
	GCAB	CABG	BACG	GBAC	BACG	EFDG	
			CBAG	GCBA	CBAG	FDEG	

Subjects in Experiment 1 were trained with 6 stimuli and tested with newly arranged ‘stems’ consisting of familiar element types. Subjects in Experiment 2 were trained with 12 stimuli and tested with new ‘stems’ consisting of novel element types. In Condition 1 the prefixation pattern predicted Go stimuli while the suffixation predicted the Nogo stimuli and vice versa for Condition 2.

Experiment 2

This experiment tested whether the zebra finches could generalize the affixation patterns to stems constructed from novel element types. It also addressed whether the discrimination shown in training and test was dependent on the presence of prefix only, suffix only, or both. A total of 12 birds were used. Six of these had previously been used in Experiment 1, another six had not been trained and tested before. The six pairs of Go and NoGo training and test stimuli in Experiment 1 were combined and used as training stimuli (Table 1). After the responses of the birds to every stimulus of the training reached criterion for at least three days, the first test started. Test 1 examined the response to new stimuli with novel stems that constructed with novel element types. The second test was given after Test 1, consisting of the ABC and DEF stems without any affix. If the birds learned both prefixation and suffixation patterns, we expected them to respond to these ‘stem-only’ stimuli at an intermediate level compared to their responses to the ‘affix-versions’. As Experiment 1, the percentages of stimulus presentations with training and test were 80:20 and the test stimuli were not reinforced.

Results

Experiment 1

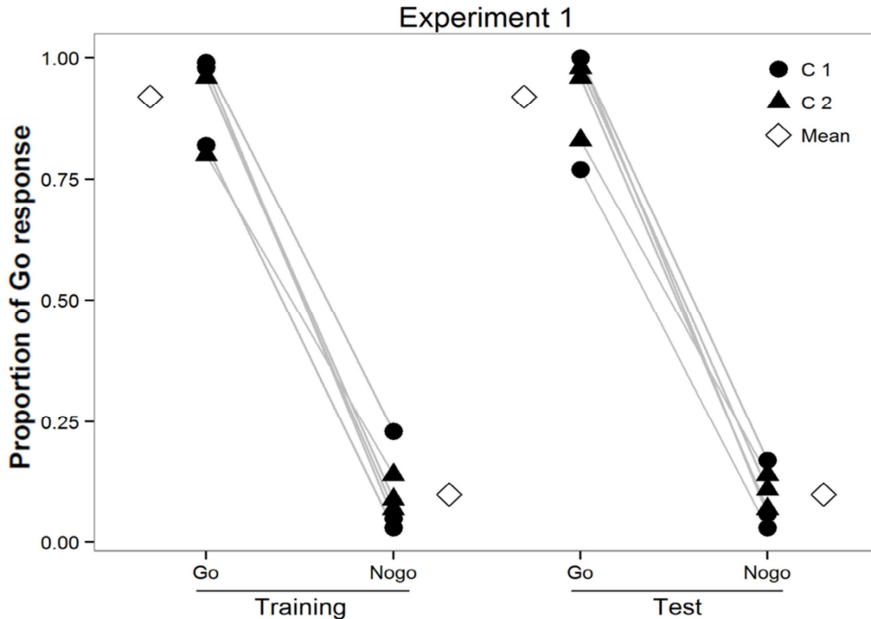


Fig. 2 Performances of individual birds of Experiment 1. All birds discriminated between prefix and suffix stimuli, both in the training and in the test, and irrespective whether the Go-stimulus was a prefix (Condition 1 (C1)) or a suffix (Condition 2 (C2)). Mean response ratios (proportion of responses to the ‘Go’ stimulus) of all birds are shown as well. Both training and test stimuli were constructed with element type A, B, and C (in different sequences), using G as affix. Test stimuli were not reinforced; ‘Go’ and ‘Nogo’ indicate test stimuli that are structurally similar to Go and Nogo training stimuli.

All birds (n=6) reached training criterion. All birds distinguished non-reinforced test stimuli with different structures as well as they discriminated the reinforced stimuli (Fig. 2). The responding ratios in the test were 0.92 ± 0.04 SEM to the Go pattern and 0.10 ± 0.02 SEM to the Nogo pattern. The Discrimination Ratio (DR, calculated as the percentage response to Go stimuli divided by the sum of the percentage response to Go stimuli and the percentage response to Nogo stimuli) for individual birds were all higher than 0.5 ($DR = 0.908 \pm 0.018$ SEM), which indicates positive discrimination (Wilcoxon Signed-rank test $Z = -2.201$, $p = 0.028$). There was no significant difference between training condition 1 and condition 2 (responses to the Go pattern: $U = 3.0$, $p = 0.700$; responses to the Nogo pattern: $U = 3.0$, $p = 0.700$, Mann-Whitney U test).

Experiment 2

Test 1

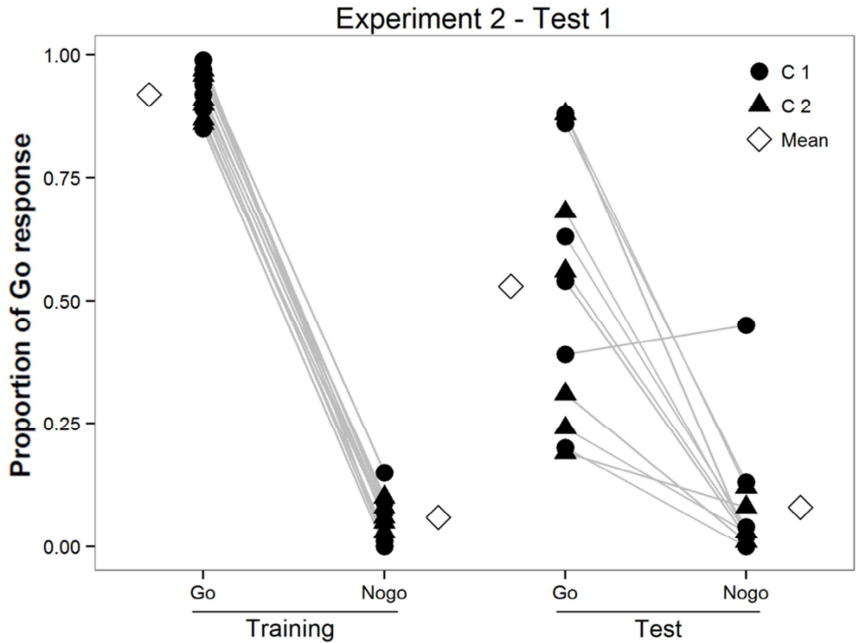


Fig. 3 Displayed are the performances of individual birds in Experiment 2, Test 1. Eleven birds discriminated between prefix and suffix stimuli in the test irrespective whether the Go-stimulus was a prefix (Condition 1 (C1)) or a suffix (Condition 2 (C2)). Mean response ratios of all birds are showed as well. Training stems were constructed with element type A, B, and C while test stems were constructed with element types D, E and F. Element G was used as the affix both in training and test. Test stimuli were not reinforced; ‘Go’ and ‘Nogo’ indicate test stimuli that are structurally similar to Go and Nogo training stimuli.

All birds (n = 12) learned to discriminate the Go and Nogo stimuli in the training. In Test 1, there was no significant difference between different training conditons (responses to the Go pattern: $U = 14.50$, $p = 0.589$; responses to the Nogo pattern: $U = 18.0$, $p = 1.0$, Mann-Whitney U test). The different training groups (Group 1 versus Group 2) also showed no significant difference (responses to the Go pattern: $U = 13.50$, $p = 0.485$; responses to the Nogo

pattern: $U = 15.50$, $p = 0.699$, Mann-Whitney U test). Therefore, the two groups were pooled. The responding ratios to the test stimuli with affixation patterns similar to Go training stimuli differed significantly from those to test stimuli with affixation patterns similar to the Nogo training stimuli (0.53 ± 0.08 SEM and 0.08 ± 0.04 SEM respectively, Fig 3. $Z = -2.934$, $p = 0.003$, Wilcoxon signed-rank). Eleven birds showed high DR in the test (0.92 ± 0.02 SEM). Only one out of 12 birds did not generalize the Go and Nogo patterns to the test stimuli (DR = 0.47).

Test 2

While the previous tests showed no differences in response patterns between the two training conditions, this test did. Therefore, data from the two training conditions are presented separately. Page's trend test for ordered alternatives (Page, 1963; Siegel & Castellan, 1981) was applied to detect whether the responses to test stimuli were ordered according to their affixes, testing the hypothesis that the responses to stimuli without an affix are expected to be in between those with a prefix or suffix

Condition 1 – Go: prefix

Responses to the stimulus without affix (ABC and DEF) were compared with their 'affix-versions' (GABC and ABCG; GDEF and DEFG). The one bird that did not generalize the Go and Nogo response to stimuli with novel element types was excluded from the test involving the DEF stem.

The test showed a significant decline in responses from GABC, ABC to ABCG ($L = 81.5$, $n = 6$, $p < 0.05$, Fig. 4a). However, most birds showed little or no differentiation between ABC and ABCG. Only one bird showed a clear intermediate response to ABC. A similar responding pattern was observed in the test with novel elements. Again a significant decline was found in responses to GDEF, DEF and DEFG ($L = 68.5$, $n = 5$, $p < 0.05$, Fig. 4b). However, the responses to the 'stem-only' stimuli were more similar to the responses to suffixed stimuli.

Condition 2 – Go: suffix

In Condition 2, the training stimuli of the Go pattern ended with a suffix while the Nogo pattern started with a prefix. The data were analyzed in the same way as those obtained for Condition 1.

The responses to stimuli decreased gradually from the Go pattern, the 'stem-only' pattern to the Nogo pattern, whether these consisted of familiar or novel element combinations. The declining trend was shown most clearly in the test with DEF stem (test with ABC stem: $L = 83$, $n = 6$, $p < 0.05$, Fig. 4c; test with DEF stem: $L = 84$, $n = 6$, $p < 0.05$, Fig. 4d).

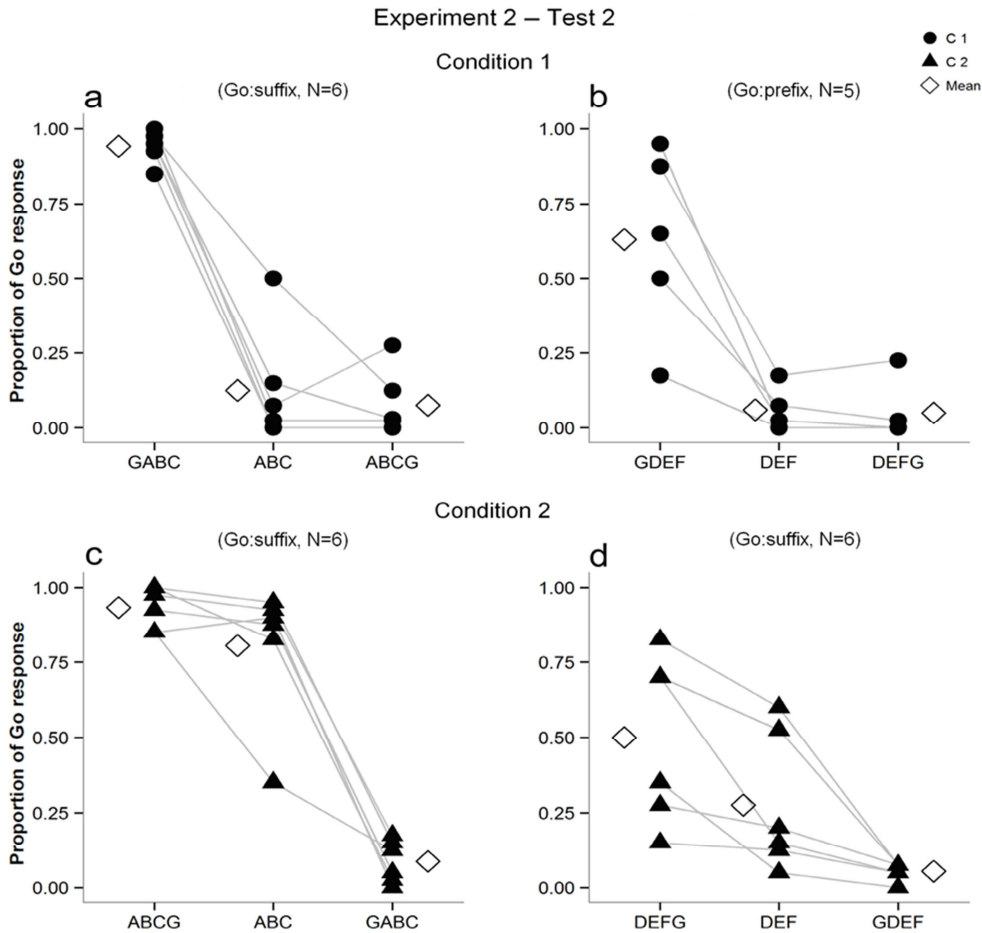


Fig. 4 Displayed are the performances of individual birds in Experiment 2, Test 2. (a) Responses to stimuli with familiar stems in Condition 1 (C1). (b) Responses to stimuli with novel stems in Condition 1. (c) Responses to stimuli with familiar stems in Condition 2 (C2). (d) Responses to stimuli with novel stems in Condition 2. Mean response ratios of all birds are showed as well.

Discussion

The results of Experiment 1 showed that the birds perfectly generalized the discrimination obtained during the training to test stimuli that shared the affixes with the training stimuli, but had a novel stem constructed from familiar element types. It shows that the discrimination resulting from the training was not based on memorization of individual stimuli, but on a shared property among Go or Nogo stimuli. This shared property could be having either a G-suffix or G-prefix. However, the result can also be obtained if the birds paid attention to whether the stimuli either started or ended with an A, B or C element. Also, if the birds use the G-element, they can achieve discrimination by attending to either the suffix position only, the prefix position only, or both. Experiment 2 addressed these questions. Test 1 shows that discrimination is maintained even when the affixations are connected to stems consisting of novel element types. So, while the somewhat lower response to test stimuli that were structurally

similar to the Go-training stimuli indicates some effect of the stem, the discrimination can only be due to presence and position of the affix: the G-element.

The results of the second test of Experiment 2 demonstrate that birds paid attention to both the prefix and suffix. However, the birds trained with suffixed sequences as Go stimuli responded to the ‘stem-only’ stimuli at a more intermediate level than the birds trained with the prefix as the Go-stimulus. This effect was less strong when the stem was ABC. Though the stimulus ABC was not affixed, it overlapped with the first part of the suffixed version used as training stimulus, and birds may have used this as an additional cue to discriminate the stimuli. The test with novel element types excluded the use of such a cue and demonstrated that at least one group attended very clearly to both prefix and suffix (Fig. 4d). The results of the two conditions together suggest that the zebra finches primarily make their judgments when listening to the beginning of the stimuli.

All birds were trained with exactly the same stimuli but the Go and Nogo associations were opposite for the two groups. Birds trained with prefixes as Go-stimuli responded strongest to stimuli starting with a ‘G’, whereas birds trained with the suffix as a Go-stimulus showed clear evidence of using both affixes. It suggests that the responses are guided both by a tendency to pay more attention to the first part of a sequence as well as paying attention to a shared feature of a stimulus set. For several songbird species there is evidence that different parts of the song may differ in their information content (e.g. Elfstrom, 1990; Kreutzer, Vallet et al., 1992; Leader, Wright et al., 2000; Lengagne, Aubin et al., 2000; Mundinger, 1975; Nelson & Poesel, 2007) and, depending on the context, either the beginning or end of songs seems most important to convey particular information. The asymmetry in attending to prefix and suffix as shown by the zebra finches may have a similar background. Interestingly, this may imply a parallel process to word recognition in human. In human linguistic studies, it has been suggested that there is a preference for suffixation in natural languages (Bybee, Pagliuca et al., 1990; Cutler, Hawkins et al., 1985; Dryer, 2005). Among the various hypotheses offered to explain the suffixation preference is the idea that suffix does not present a problem for making word recognition more difficult while prefix does (Dryer, 2005). It has been proposed that the beginning of a word is its most psychologically salient part (Clark, 1991) and it is important for spoken word activation (Marslen-Wilson, 1987; Rodd, 2004). In the prefixed word, the processing of the stem does not precede the affix, so it is not easy to do an on-line exploitation of the information of the whole word (Kandel, Spinelli et al., 2012). The suffixation preference in language is proposed to be driven by a cognitive mechanism that is not specific to language (Hupp, Sloutsky et al., 2009) and our results suggests it may even be by a mechanism not specific to humans.

So, we conclude that our results provide the first evidence in a non-primate of learning a rule that, at least in its surface pattern, is similar to a linguistic affixation pattern: birds can identify that the presence of a specific vocal unit at one or the other edge of a string is linked to different ‘meanings’ and can generalize this knowledge to novel strings. As songbirds are phylogenetically quite distant from humans, our findings are unlikely to indicate a shared ancestral trait but might be the result of independent evolution. Also for this reason, the similarity should not be taken as evidence that the competence is formally fully similar to that of

humans using affixations. As outlined by Berwick et al. (2011), there is quite a gap between the syntactic structures birds use or can detect and those present in human languages. However, our experiment indicates a processing competence that may also have been present in pre-linguistic humans and acted as a domain-general perceptual and memory primitive (Endress et al. 2009) that has been co-opted for the evolution of a linguistic phenomenon.

References

- Berwick, R. C., Okanoya, K., Beckers, G. J., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in cognitive sciences*, 15(3), 113-121. doi: 10.1016/j.tics.2011.01.002
- Bolhuis, J. J., & Everaert, M. (2013). *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain*. Cambridge: MIT Press.
- Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, 11(11), 747-759. doi: 10.1038/Nm2931
- Bybee, J. L., Pagliuca, W., & Perkins, R. D. (1990). On the asymmetries in the affixation of grammatical material. In W. Croft, K. Denning & S. Kemmer (Eds.), *Studies in typology and diachrony*. Amsterdam: John Benjamins Publishing Company.
- Clark, E. V. (1991). Acquisitional principles in lexical development. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development*. Cambridge, MA: Cambridge University Press.
- Comins, J. A., & Gentner, T. Q. (2010). Working memory for patterned sequences of auditory objects in a songbird. *Cognition*, 117(1), 38-53. doi: 10.1016/j.cognition.2010.06.009
- Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The Suffixing Preference, a Processing Explanation. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, 23(5), 723-758. doi: DOI 10.1515/ling.1985.23.5.723
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. doi: 10.1146/annurev.neuro.22.1.567
- Dryer, M. S. (2005). Prefixing versus suffixing in inflectional morphology. In M. Haspelmath, M. S. Dryer, D. Gil & B. Comrie (Eds.), *The world atlas of language structures*. New York: Oxford University Press.
- Elfstrom, S. T. (1990). Responses of Territorial Meadow Pipits to Strange and Familiar Song Phrases in Playback Experiments. *Animal Behaviour*, 40, 786-788. doi: Doi 10.1016/S0003-3472(05)80712-2
- Endress, A. D., Cahill, D., Block, S., Watumull, J., & Hauser, M. D. (2009). Evidence of an evolutionary precursor to human language affixation in a non-human primate. *Biology Letters*, 5(6), 749-751. doi: 10.1098/rsbl.2009.0445
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal cognition*, 13(3), 483-495. doi: 10.1007/s10071-009-0299-8
- Endress, A. D., & Hauser, M. D. (2011). The influence of type and token frequency on the acquisition of affixation patterns: implications for language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 77-95. doi: 10.1037/a0020210
- Endress, A. D., & Mehler, J. (2010). Perceptual constraints in phonotactic learning. *Journal of Experimental Psychology: Human Perception and Performance*, 36(1), 235-250. doi: 10.1037/a0017164
- Endress, A. D., Nespore, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in cognitive sciences*, 13(8), 348-353. doi: 10.1016/j.tics.2009.05.005
- Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, 134(3), 406-419. doi: 10.1037/0096-3445.134.3.406
- Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088), 1204-1207. doi: 10.1038/nature04675
- 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(5598), 1569-1579. doi: 10.1126/science.298.5598.1569
- Henson, R. N. (1998). Short-term memory for serial order: the Start-End Model. *Cognitive Psychology*, 36(2), 73-137. doi: 10.1006/cogp.1998.0685
- Henson, R. N. (1999). Positional information in short-term memory: relative or absolute? *Memory & Cognition*, 27(5), 915-927.
- Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly journal of experimental psychology. A, Human experimental psychology*, 49(1), 116-139. doi: Doi 10.1080/027249896392829
- Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (2009). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes*, 24(6), 876-909. doi: 10.1080/01690960902719267
- Kandel, S., Spinelli, E., Tremblay, A., Guerassimovitch, H., & Alvarez, C. J. (2012). Processing prefixes and suffixes in handwriting production. *Acta Psychologica*, 140(3), 187-195. doi: 10.1016/j.actpsy.2012.04.005
- Kreutzer, M., Vallet, E., & Doucet, S. (1992). Sexual-Responses of Female Canaries at the Onset of Song Stimuli. *Experientia*, 48(7), 679-682. doi: Doi 10.1007/Bf02118318
- Kuczaj, S. A. (1977). Acquisition of Regular and Irregular Past Tense Forms. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 589-600. doi: Doi 10.1016/S0022-5371(77)80021-2

- Leader, N., Wright, J., & Yom-Tov, Y. (2000). Microgeographic song dialects in the orange-tufted sunbird (*Nectarinia osea*). *Behaviour*, 137, 1613-1627. doi: Doi 10.1163/156853900502745
- Lengagne, T., Aubin, T., Jouventin, P., & Lauga, J. (2000). Perceptual salience of individually distinctive features in the calls of adult king penguins. *The Journal of the Acoustical Society of America*, 107(1), 508-516.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Marslen-Wilson, W. D., & Tyler, L. K. (1998). Trends in cognitive sciences. *Trends in cognitive sciences*, 2(11), 428-435.
- Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: an exploratory study. *System*, 28(2), 291-304. doi: dx.doi.org/10.1016/S0346-251X(00)00013-0
- Mundinger, P. (1975). Song Dialects and Colonization in House Finch, *Carpodacus-Mexicanus*, on East Coast. *Condor*, 77(4), 407-422. doi: Doi 10.2307/1366088
- Nagy, W. E., Diakidoy, I. A. N., & Anderson, R. C. (1993). The Acquisition of Morphology - Learning the Contribution of Suffixes to the Meanings of Derivatives. *Journal of Reading Behavior*, 25(2), 155-170.
- Nelson, D. A., & Poesel, A. (2007). Segregation of information in a complex acoustic signal: individual and dialect identity in white-crowned sparrow song. *Animal Behaviour*, 74, 1073-1084. doi: DOI 10.1016/j.anbehav.2007.01.018
- Ohms, V. R., Gill, A., Van Heijningen, C. A., Beckers, G. J., & ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 277(1684), 1003-1009. doi: 10.1098/rspb.2009.1788
- Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, 404(6773), 77-80. doi: 10.1038/35003571
- Page, E. B. (1963). Ordered hypotheses for multiple treatments: a significance test for linear ranks. *Journal of the American Statistical Association*, 58(301), 216-230.
- Rodd, J. M. (2004). When do leotards get their spots? Semantic activation of lexical neighbors in visual word recognition. *Psychonomic bulletin & review*, 11(3), 434-439.
- Seki, Y., Suzuki, K., Osawa, A. M., & Okanoya, K. (2013). Songbirds and humans apply different strategies in a sound sequence discrimination task. *Frontiers in psychology*, 4, 447. doi: 10.3389/fpsyg.2013.00447
- Siegel, S., & Castellan, J. N. (1981). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- St Clair, M. C., Monaghan, P., & Ramsar, M. (2009). Relationships between language structure and language learning: the suffixing preference and grammatical categorization. *Cognitive science*, 33(7), 1317-1329. doi: 10.1111/j.1551-6709.2009.01065.x
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1984-1994. doi: 10.1098/rstb.2012.0055
- Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, 14(1), 66-73.
- 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(2), 165-175. doi: 10.1007/s10071-012-0559-x
- 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. *Proceedings of the National Academy of Sciences*, 106(48), 20538-20543. doi: 10.1073/pnas.0908113106
- Wright, A. A., Santiago, H. C., Sands, S. F., Kendrick, D. F., & Cook, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, 229(4710), 287-289.

Chapter 4^{*}

Artificial grammar learning in zebra finches and human adults: XYX vs XXY

^{*} This chapter is based on: Jiani Chen, Danielle van Rossum & Carel ten Cate (in press). Artificial grammar learning in zebra finches and human adults: XYX vs XXY . *Animal Cognition*

Abstract

Abstracting syntactic rules is critical to human language learning. It is debated whether this ability, already present in young infants, is human- and language-specific or can also be found in non-human animals, indicating it may arise from more general cognitive mechanisms. Current studies are often ambiguous and few have directly compared rule learning by humans and non-human animals. In a series of discrimination experiments, we presented zebra finches and human adults with comparable training and tests with the same artificial stimuli consisting of XYX and XXY structures, in which X and Y were zebra finches song elements. Zebra finches readily discriminated the training stimuli. Some birds also discriminated novel stimuli when these were composed of familiar element types, but none of the birds generalized the discrimination to novel element types. We conclude that zebra finches show evidence of simple rule abstraction related to positional learning, suggesting stimulus bound generalization, but found no evidence for a more abstract rule generalization. This differed from the human adults, who categorized novel stimuli consisting of novel element types into different groups according to their structure. The limited abilities for rule abstraction in zebra finches, may indicate what the precursors of more complex abstraction as found in humans may have been like.

Introduction

Humans exposed to a string of meaningless speech items (e.g. CV-syllables) arranged according to a specific algorithm, are generally well able to detect the underlying rule (Reber, 1967, 1990). Such syntactic rule abstracting is considered a hallmark of human linguistic abilities (Hauser, Chomsky et al., 2002; Marcus, Vijayan et al., 1999; Pinker, 1991). However, the ability for rule abstraction is also shown when the items are non-speech stimuli, like musical tones and other non-linguistic sounds (Endress, 2010; Gebhart, Newport et al., 2009; Saffran, Johnson et al., 1999, but see Creel, Newport et al., 2004) or pictures (Saffran, Pollak et al., 2007). Findings like this raise the question whether rule abstraction mechanisms are specific to language or more domain general (Folia, Udden et al., 2010; Gomez & Gerken, 2000). They also raise the question whether the ability for rule abstraction from structured acoustic input is specific to humans or shared with other species, providing a general cognitive basis for the evolution of human syntactical abilities. Various animal species have demonstrated an ability for abstraction in different contexts, such as for the concept of same vs different, which has been demonstrated in species ranging from bees (Giurfa, Zhang et al., 2001) to pigeons (Katz & Wright, 2006) and monkeys (Wright, Rivera et al., 2003). Against this background several studies have addressed the abilities of various non-human animals to detect and generalize regularities in acoustic string structures. Such a task requires the animals to first detect the structure of individual strings and next to notice the abstract similarity between different strings. In this respect it is similar to discrimination learning based on second order relationships, a task that chimpanzees, but not monkeys, can do using visual stimuli (Thompson & Oden, 2000).

Like humans, animals can be exposed to strings of meaningless sound and next be tested with strings consisting of novel items and/or novel arrangements of items. The responses to such novel stimuli allow conclusions about whether the subjects have been able to detect the underlying algorithm. Various experiments have used this paradigm to examine rule detection in primates (Endress, Carden et al., 2010; Fitch & Hauser, 2004; Hauser & Glynn, 2009; Hauser, Newport et al., 2001; Newport, Hauser et al., 2004); rats (de la Mora & Toro, 2013; Murphy, Mondragon et al., 2008; Toro & Trobalon, 2005); and birds (Gentner, Fenn et al., 2006; Herbranson & Shimp, 2003, 2008; Stobbe, Westphal-Fitch et al., 2012; Seki, Suzuki et al., 2013; van Heijningen, Chen et al., 2013; van Heijningen, de Visser et al., 2009; Comins & Gentner, 2014; Comins & Gentner, 2013; Yamazaki, Suzuki et al., 2012). However, the degree to which animals can detect and generalize rules to novel items – the hallmark of rule abstraction – is by no means clear, as the results of several studies can be explained by simpler mechanisms such as, e.g., phonetic generalization to novel strings or items (Beckers, Bolhuis et al., 2012; Corballis, 2007, 2009; Fitch & Friederici, 2012; Seki, Suzuki et al., 2013; ten Cate & Okanoya, 2012; van Heijningen, de Visser et al., 2009). Hence, there is a need for comparative studies, addressing similarities and differences in how humans and animals cope with extracting, processing and generalizing specific types of input.

In the current study we compare the abilities of zebra finches and human adults to detect the difference between an *XYX* and a *XXY* structure, where X and Y denote arbitrary items. In a classic study, Marcus, Vijayan et al. (1999) showed that 7-month old infants

habituated to *XXY* or *XYX* speech stimuli (such as ‘gatiga’ or ‘gatiti’, X and Y being different syllables) could extract the underlying regularity and apply it to novel stimuli composed of new syllables. The infants could also distinguish *XXY* from *XYX* strings. A subsequent study showed even 5-month old infants to have such ability, provided that they are trained and tested with congruent combinations of vocal and visual stimuli (Frank, Slemmer et al., 2009). Although the mechanisms that underlie the infants’ ability for generalization evoked much discussion (e.g. Altmann & Dienes, 1999; Eimas, 1999; Marcus, Vijayan et al., 1999; Negishi, 1999; Seidenberg, 1999), the observation that the infants are able to discriminate between stimuli with different structures and generalizing this to novel items without basing on the physical sameness of items is broadly accepted. As regards the abilities of non-human animals to detect and generalize such structures, the results are ambiguous. The first study to address this (Hauser, Weiss et al., 2002) suggested that tamarins could resolve the task, but this study has been retracted. Later on, Hauser and Glynn (2009) suggested that rhesus monkeys could discriminate *XYX* from *XXY* when X and Y were rhesus monkey calls. However, the strings used to test for rule learning were composed of the same call types as the ones used for training, i.e. they concerned within-category generalization, which, as demonstrated in other studies (e.g. Gentner et al. 2006; van Heijningen et al. 2009) several species are able to do. However, such findings leave the question for between-category generalization of the rule to truly novel items unanswered. Other studies have examined whether rats could extract and distinguish the grammar rules. Toro and Trolldenier (2005) failed to find evidence for discriminating *XXY* from *XYX*, with X and Y being human speech stimuli, although recently de la Mora and Toro (2013) showed rats to be able to distinguish *XXY* from *XYZ* sequences. Using a different experimental design Murphy, Mondragon et al. (2008) claimed that rats could distinguish *XYX* from *XYX* and *XXY*. However, since the stimuli of this experiment were composed by using pure tones, it has been questioned whether this transfer is really based on extracting the underlying abstract rule by the rats or whether they generalized based on the tonal contour (Corballis, 2009; but see Mondragon, Murphy et al., 2009). So, the studies on mammals do not yet provide a conclusive answer to whether animals can abstract the rules underlying similarly structured auditory strings.

In our study we use a songbird, the zebra finch, to address this question. Songbirds are phylogenetically very distant from humans. Yet, their natural songs show many similarities to speech and language, making them of special interest for examining rule learning abilities. In contrast to many other species, including primates, songbirds have complex, learned, vocalizations consisting of strings of phonetically different elements (‘notes’). While the acoustic features of these elements may vary between individuals of a species, different individuals structure their element strings according to species specific sequencing rules. In addition, birdsong researchers have discovered remarkable neural and genetic parallels between the processes involved in the production, perception, and development of language and birdsong (e.g. Berwick, Okanoya et al., 2011; Bolhuis, Okanoya et al., 2010; Doupe & Kuhl, 1999; Ohms, Escudero et al., 2012). These similarities raise the question whether songbirds also share more advanced cognitive abilities involved in language processing, in particular in the area of rule abstraction. Establishing to what extent the rule abstraction capabilities in songbirds match (or differ from) those of humans will help to understand the relation between these capabilities and

the presence and evolution of vocal complexity, even though the similarities might be based on parallel evolution rather than reflecting a shared ancestral state.

Some studies have addressed rule abstraction in several songbird species, but, as in mammals, have so far not provided clear answers to the question of their ability for such abstraction (reviewed by ten Cate & Okanoya, 2012). With respect to the ability of songbirds to distinguish XYX from XXY and XYY strings, a first study was done by van Heijningen, Chen et al. (2013). In that study we examined the ability of zebra finches to discriminate such strings, using a ‘Go-NoGo’ design. We showed that zebra finches could discriminate two ‘Go’ stimuli with an XYX structure (ABA and BAB, with A and B indicating two different song elements) from four ‘NoGo’ stimuli (AAB, BBA, ABB and BAA). Using non-reinforced probe sounds having different structures, we demonstrated that the distinction was most likely based on detecting whether or not the stimulus strings contained AA or BB repetitions, rather than using a more complex abstract rule such as comparing the first and last items of strings, as suggested by Marcus, Vijayan et al. (1999). Even though the generalization to novel items was very limited (present in one bird out of eight), this finding indicates a sensitivity to repetition that is also present in humans (Endress, Dehaene-Lambertz et al., 2007; Gervain, Macagno et al., 2008; Kovács & Mehler, 2009) and may hint at a primitive rule-like regularity that may constitute a basis for the evolution of a higher level of abstraction (Endress, Nespor et al., 2009). However, the limited evidence for abstraction in our experiment might be due to the training in which only A and B elements were used to create XYX, XXY and XYY strings, precluding the birds from forming a more abstract representation of the string structure. Hence the results leave open the question whether any non-human animals can detect an abstract rule from training with a more extensive set of examples. In the current study we trained zebra finches to discriminate larger sets of XYX from XXY strings and examined whether they generalized the distinction to strings consisting of novel items (Experiments 1a and 1b).

As the operant paradigm differs from the habituation paradigm as used in the studies in human infants and monkeys, which has advantages and disadvantages (ten Cate & Okanoya, 2012), we also tested human adults with the same stimuli in a comparable operant design as used for the zebra finches (Experiments 2a and 2b) to allow a more direct comparison of zebra finches and humans.

Rule learning in zebra finches - Experiment 1a

Material and Method

Subjects and apparatus

Six zebra finches (3 males and 3 females), which were all naïve to any experimental setup, were individually trained and tested in operant conditioning cages [70 (l) × 30 (d) × 45 (h) cm] in sound attenuated chambers using a Go/No-go paradigm. Each cage was made of wire mesh with a plywood back wall. The back wall contained a food hatch and two red pecking keys, each containing a red LED. A fluorescent tube on the top of the cage emitted daylight spectrum light on a 13.5 L: 10.5 D schedule (except for a penalty during which the light was switched off temporarily). In the training and test periods, the first pecking key was illuminated

with a red LED. A peck on the first key resulted in a sound stimulus and also activated the second key, indicated by switching on the LED of this key. Depending on the sound stimulus, a peck on the second key resulted in a food reward or a mild punishment of darkness (see procedure below; for more detailed description about the apparatus, see van Heijningen, Chen et al., 2013).

Stimuli

Stimuli were constructed from zebra finch song elements originating from our zebra finch song database, consisting of undirected song. Natural zebra finch song consists of a series of 4-12 elements, which are sung in a relatively fixed sequence. These elements can be considered the natural units of production and perception in zebra finch songs (ten Cate, Lachlan et al., 2013). At the species level, song elements can be categorized into around eight different types and there are no strong constraints to the position of each element within the song (Lachlan, Verhagen et al., 2010). There is no indication that single elements or element combination carry any meaning and hence these elements seem suited to construct artificial meaningless strings. Based on optimal discriminability, six element types were selected ('flat,' 'trill,' 'downslide,' 'high,' 'curve,' and 'noisy', see Fig. 1). They were ramped and the RMS values of different song elements were equalized in PRAAT (version 5.1.15, www.praat.org).

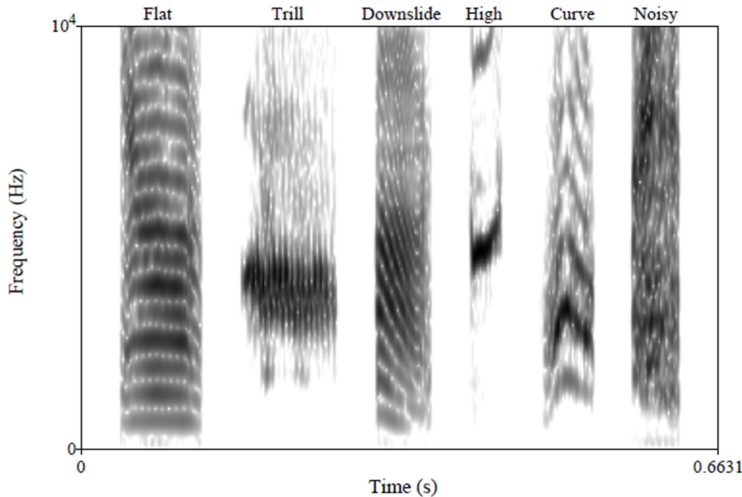


Fig. 1 Spectrograms of the six element types.

Stimuli were constructed following *XYX* and *XXY* patterns. The *XYX* corpus contained triplets such as ABA, BCB, CDC, DAD, EFE etc. (A, B, C, D, E, F representing specific different song elements), and the *XXY* corpus contained triplets such as AAB, BBC, CCD, DDA, EEF etc (Table 1). Elements A, B, C and D were used to compose training stimuli. Another two elements E and F served as test elements and were not presented in the training phases. *XYX* stimuli were always presented as Go stimuli and *XXY* as Nogo stimuli during

training (in an earlier experiment, using a somewhat different grammar we obtained no evidence that it mattered which stimulus was the Go and which the Nogo, van Heijningen, de Visser et al., 2009). Forty ms of silence was inserted between consecutive elements of the various element sequences, which is comparable to the pauses that separate song elements in natural songs. To eliminate pseudo effect of an arbitrary parameter of the sounds, the element assignment was shuffled for the subjects, for instance, an ABA triplet could be ‘flat-downslide-flat’ for one bird and ‘high-trill-high’ for another bird.

Table 1 Training and test stimuli of Experiment 1a

Phase	Training stimuli		control	Test stimuli (non-reinforced)		
	Go (XYX)	NoGo (XXY)		XYX	XXY	XXY
1	ABA	AAB	ABA	BAB	BBA	
	ACA	AAC	AAB	BCB	BBC	
	ADA	AAD		BDB	BBD	
2	BAB	BBA	BAB	CAC	CCA	
	BCB	BBC	BBA	CBC	CCB	
	BDB	BBD		CDC	CCD	
3	CAC	CCA	CAC	DAD	DDA	
	CBC	CCB	CCA	DBD	DDB	
	CDC	CCD		DCD	DDC	
4	DAD	DDA		ABA	AAB	
	DBD	DDB	DAD	ADA	AAD	
	DCD	DDC	DDA	BCB	BBC	
				CAC	CCA	
			CDC	CCD		
5	ABA	AAB				
	ADA	AAD				
	BCB	BBC	ABA	EFE	EEF	FEE
	CAC	CCA	AAB	FEF	FFE	EFF
	CDC	CCD				
	DAD	DDA				

Phases 1–4 consist of 6 training stimuli and phase 5 consists of 12 training stimuli, selected from the 4 preceding training phases. Each letter indicates a different song element. Control test stimuli were identical to the training stimuli but not reinforced. The other test stimuli were different from training stimuli and not reinforced either.

Procedure

The zebra finches were first trained to discriminate a natural zebra finch song (Go stimulus) from a pure tone (Nogo stimulus). Pecking on the first key initiated a playback of either a song or a tone. The birds gained access to food for 10s as reward when they pecked the second key after hearing a Go stimulus. Inversely, when they pecked the second key after hearing a Nogo stimulus, they received a mild punishment with the light extinguishing for 15s.

After the responses of the birds to the song and tone reached training criterion (>75% response to Go stimuli and <25% response to Nogo stimuli) for at least two full consecutive days, the training began, using the XYX stimuli as Go and the XXY stimuli as NoGo stimuli. This experiment contained 5 successive phases, each of which used a different set of training and test stimuli (Table 1).

When the birds reached training criterion for at least three full consecutive days, they entered a test phase. In test phases, test stimuli were not reinforced to avoid additional learning. They were presented in only 20% of all stimulus presentations to prevent extinction of the pecking behavior. The other 80% of stimulus presentations consisted of the reinforced stimuli used during the training phases. Every test contained 40 presentations of each test stimulus. Control test stimuli identical to the training stimuli (but not reinforced) were added to control for the effect of non-reinforcement. After finishing the test of phase 1, the test stimuli became the training stimuli for the next training phase. Testing for this phase started when the learning criterion had been reached again. When the birds finished the 4th training, they were tested with 6 pairs of stimuli (including control stimuli) selected from the 4 preceding training phases (see Table 1) to address whether they combined the information of the preceding phases. We next trained the birds with these 6 pairs of stimuli to enhance the learning of XYX and XXY rules. After this, a final test consisted of stimuli composed of new element types (E and F, Table 1). This test included an additional XYY structures (EFF and FEE) to examine whether this was discriminated from the XYX and XXY structures.

Some birds had a persistent lower response proportions to one or more stimuli during the final training phase. However, if the overall Discrimination Ratio (DR, calculated as the percentage response to Go stimuli divided by the sum of the percentage response to Go stimuli and the percentage response to Nogo stimuli) was > 0.75 for at least three consecutive days, the bird entered the test phase.

Performance evaluation

To examine whether the birds transferred the discrimination learned during the training to the test stimuli, we calculated the DR for the XYX and XXY test stimuli. A DR higher than 0.50 indicated positive discrimination. We used Wilcoxon signed ranks test to examine the differences in responses to XYX and XXY strings. Simultaneous Testing Procedures (STP) based on G-tests were applied to detect how the birds categorized different test stimuli and to examine the performances of individual birds (Gabriel, 1966; Sokal & Rohlf, 1995; A Williams correction was calculated when numbers were lower than 5; G value $> X^2$ critical value means $P < 0.05$ and vice versa).

Results

All birds reached criterion during training, but the training trials varied between different training phases. From training 1 to 5, birds took 3418 ± 692.47 , 3262 ± 594.77 , 4066 ± 1071.71 , 2345 ± 431.88 and 3895 ± 359.80 (Mean $\pm$ SEM, $n = 6$) trials respectively to reach training criterion.

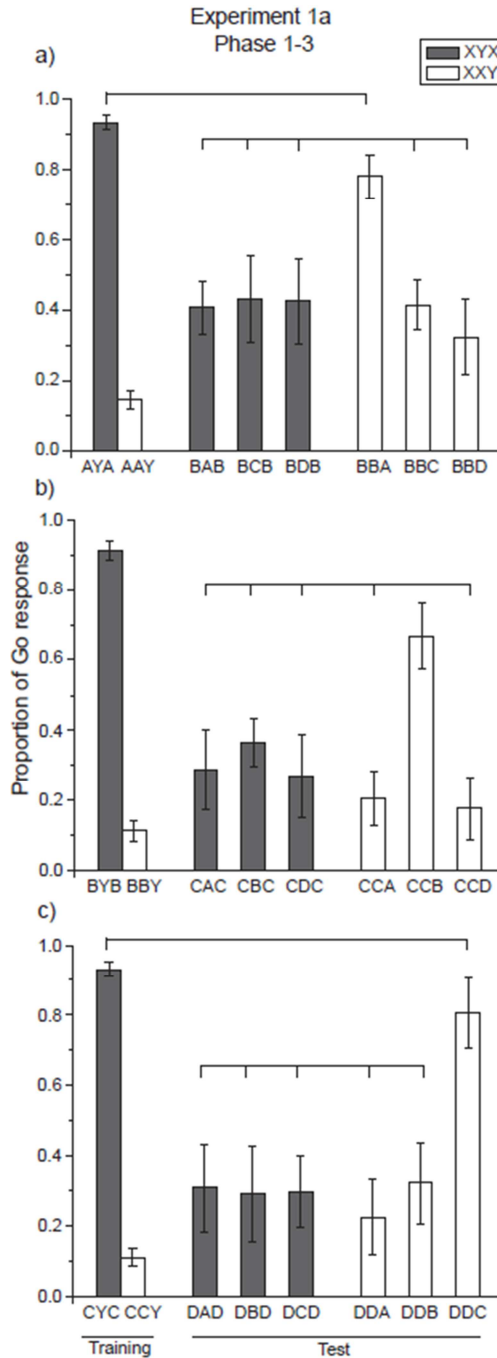


Fig. 2 Mean ($\pm$ SEM, $n = 6$) proportion of Go responses to the training and test stimuli in the a) first b) second and c) third phase of Experiment 1a. Y in the training stimuli refers to the three elements that can occupy this position, e.g. AYA includes ABA, ACA and ADA (see Table 1). The bars in the Test –section refer to the proportion of ‘Go’ responses to the different test stimuli. Lines above the bars indicate homogenous responses (no significant differences, X^2 critical value of 14.07) to different stimuli sets. Test stimuli that end with A in test phase 1 (Fig. 2a), B in test phase 2 (Fig. 2b) and C in test phase 3 (Fig. 2c) get more responses in comparison with the other test stimuli in the same test phase.

During all the tests, the responses to the positive control stimuli and the positive training sets showed no significant differences nor did the responses to the negative control stimuli and their negative training sets (all $P > 0.05$). This indicates that the non-reinforcement condition did not have an impact on the performances of the birds. Therefore, we compared the responses to the test stimuli with the average responses to the positive and the negative training stimuli.

The results of test phase 1 showed that proportions of responses to BAB, BCB, BDB, BBC and BBD were all between the proportions of the Go and Nogo stimuli (XYX and XXY stimuli of training). They were all in the same homogenous group (no significant difference from each other, $G = 1.20 < X^2$ critical value 14.07, $df = 4$, Fig 2a). The exception is provided by test stimulus ending with A (BBA), which is being treated as a Go stimulus (ABA, ACA and ADA, $G = 10.19$, $df = 1$, X^2 critical value = 14.07, Fig 2a). This pattern is repeated in the test phases after the second and third training, where again all test stimuli form a homogenous group (all $G < X^2$ critical value 14.07, $df = 4$) while there is one stimulus (ending with B: CCB and ending with C: DDC, in test 2 and 3 respectively) with a score close to the Go stimuli (stimuli BAB, BCB and BDB in training 2 and stimuli CAC, CBC and CDC in training 3, both $G < X^2$ critical value 14.07, $df = 1$, Fig 2b,c).

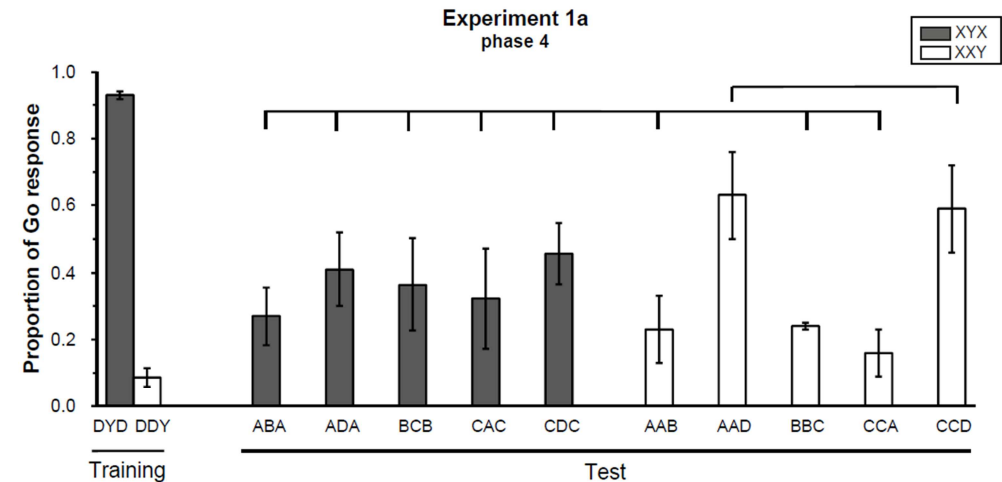


Fig. 3 Mean ($\pm$ SEM, $n = 6$) proportion of Go responses to the training and test stimuli in phase 4 of Experiment 1a. Test stimuli consisted of familiar elements known from the training. Y in the training stimuli refers to the elements A, B and C, for which the values were averaged. Lines above the bars indicate homogenous responses (no significant differences, X^2 critical value of 19.68) to different stimuli sets.

In test 4, the responses to different test stimuli were not based on XYX and XXY structures (responses to ABA, ADA, BCB, CAC, CDC, AAB, BBC and CCA were all in the same homogeneous group, $G = 13.56 < X^2$ critical value 19.68, $df = 7$, Fig 3), which indicates that the birds did not combine what they had learned in the 4 preceding training phases. Among the responses to the separate test stimuli, responses to AAD and CCD were higher, to which,

like the positive training set, had a D element in the final position (no significant difference from each other, $G = 0.21$, X^2 critical value = 19.68, $df = 1$).

A final test (Test 5), this time with song elements E and F (not used in any of the previous stimuli) was given after a training with the 6 stimulus pairs from the 4th test (Table 1). Again the responses to the *XYX* and *XXY* test stimuli were not significantly different ($Z = -1.214$, $P = 0.225$, $n=6$). In addition, the responses to the test stimuli of the *XYX*, *XXY* and *XXY* structures were all in the same homogenous group ($G = 0.24 < X^2$ critical value 9.49, $df = 2$), which was significantly different from the responses to the training stimuli (Fig 4).

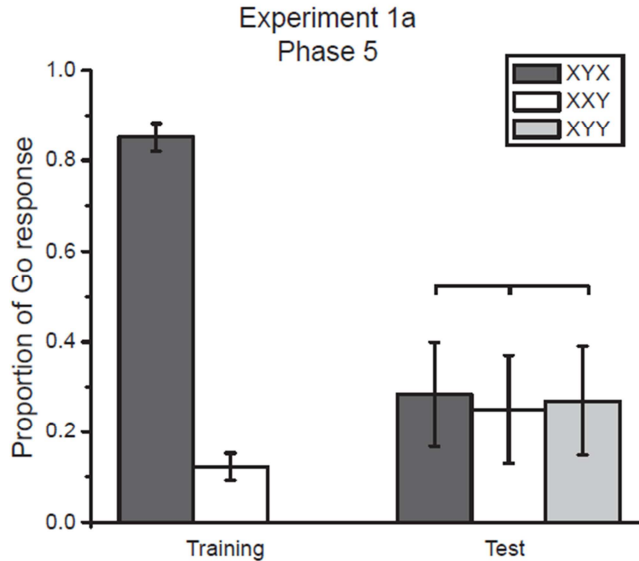


Fig. 4 Mean ($\pm$ SEM, $n = 6$) proportion of Go responses to training and test stimuli in phase 5 of Experiment 1a. Test stimuli consisted of novel elements not present in training. Lines above the bars indicate homogenous responses (no significant differences, X^2 critical value of 9.49) to different stimuli sets.

Discussion

The results of the training demonstrate that zebra finches are able to discriminate strings in which the same elements are arranged in different sequences. However, the results of the first four tests suggest that the birds did not use the underlying *XYX* and *XXY* structures of the strings to guide their discrimination. Nevertheless, they differentiated among the test strings in a systematic way. The deviant test stimuli that were being treated more similar to *XYX* training stimuli in tests 1, 2, 3 and 4 were always the test stimuli that shared their final element with the *XYX* training stimuli (test 1: BBA - element A; test 2: CCB - element B; test 3: DDC - element C; test 4: AAD, CCD – element D). The most likely explanation is that the birds learned to use the final elements of the *XYX* strings to discriminate ‘good’ and ‘bad’ sounds, demonstrating a stimulus bound generalization. Attending to this final element can also explain the other results of test 4. Although the set contained stimuli that were very well discriminated in

earlier phases, in each subsequent training phase the birds had to relearn which final element was associated with obtaining the food reward. E.g. the final A element that indicated a 'Go' stimulus in phase 1 became a 'NoGo' stimulus in training phase 2, which is likely to have been at the expense of the earlier discrimination. To prevent the birds from using this regularity, we gave them the extended training set in phase 5. Every position of the song sequences of the training stimuli could be occupied by any of the four element types (A, B, C and D, see Table 1). Therefore the birds now had to use another strategy to discriminate between the positive and negative stimuli, such as paying attention to the abstract structure of the strings or rote memorization of individual stimuli. If the birds used abstract rules to make the distinction they were expected to respond to novel songs with novel element types according to the overall structures of the stimuli. In addition, if they could detect the XYX and XXY rules they might also notice that a XYY rule was different. However, the results of test 5 showed no difference in responding to the three stimulus structures. We conclude that the birds failed to detect the underlying structures of the stimuli and most likely relied on having learned particular elements in relation to their position (phase 1-4), or memorized the individual training strings (phase 5) to discriminate the training sets.

Rule learning in zebra finches - Experiment 1b

It is possible that the training procedure of Experiment 1a biased the birds' learning process away from rule learning and towards tracking final elements. This may also have hampered rule learning in phase 5. Experiment 1b aimed at avoiding such a bias. To this end, a new group of naïve zebra finches was trained by immediately starting with the stimuli used in phase 5 of the previous experiment.

Material and Method

Subjects and apparatus

Eight zebra finches (4 males and 4 females) were trained and tested in this experiment. The experimental conditions were identical to phase 5 of Experiment 1a.

Stimuli

Stimuli from the fifth training phase of the Experiment 1a were used as training stimuli in this experiment. After training, birds were switched to two successive test phases. Similar to test phase 5 of Experiment 1a, the first test contained novel stimuli with novel element types; for the second test the stimuli were constructed by re-arranging familiar element types (Table 2). This test was added to explore whether the discrimination shown in the training was maintained when familiar elements were combined in novel combinations.

Procedure

The procedure and criteria of performance were identical to Experiment 1a.

Performance evaluation

The statistical methods were the same as for Experiment 1a.

Table 2 Training and test stimuli of Experiment 1b

Training stimuli		Test stimuli (non-reinforced)			
Go (XYX)	NoGo (XXY)	control	XYX	XXY	XXY
ABA	AAB	ABA AAB	EFE	EEF	FEE
ADA	AAD		FEF	FFE	EFF
BCB	BBC				
CAC	CCA		BAB	BBA	
CDC	CCD		CBC	CCB	
DAD	DDA	2	ACA	AAC	
			DCD	DDC	

Letters indicate different song elements. See Legend of Table 1 for further explanation

Results

Seven out of eight birds learned to discriminate between the two sets of the training stimuli (10494 ± 1289.96 Mean $\pm$ SEM trials were required until they reached training criterion, $n = 7$). The responses to the control stimuli and their training sets again showed no significant differences (all $p > 0.05$), which indicates that here also the non-reinforcement condition had no impact on the responses.

The results of Test 1 of this experiment showed that birds did not discriminate between the *XYX*, *XXY* and *XXY* test stimuli constructed from the novel elements E and F. The responses to all three test stimuli were in the same homogeneous group with those to negative training stimuli ($G = 0.21$, $df = 3$, X^2 critical value = 12.59) and significantly different from those to the positive training stimuli ($G = 688.35$, $df = 4$, X^2 critical value = 12.59, Fig 5). However in the second test, with familiar items in novel positions, the responses to *XYX* and *XXY* were significantly different ($G = 18.56$, $df = 1$, X^2 critical value = 12.59). The responses to *XYX* stimuli in Test 2 were similar to *XYX* stimuli in training ($G = 1.23$, $df = 1$, X^2 critical value = 12.59), whereas the responses to *XXY* structure in Test 2 were significantly different from those to other structures (all $G > 12.59$). When analysed at the individual level, the DRs between these two sets of stimuli were high in two birds (0.88 and 0.70). The responses of these two birds to the *XYX* test stimuli were significantly different from the responses to the *XXY* test stimuli (all $G > X^2$ critical value of 7.851, see Fig 6a for an example). Another two birds had lower DRs of 0.58 and 0.54, but the responses to the *XYX* and the *XXY* test stimuli also differed significantly (all $G > X^2$ critical value of 7.851). Finally, three birds showed no clear patterns in discriminating between the test stimuli (DRs = 0.50 ± 0.01 SEM, $n = 3$, all $G < X^2$ critical value of 7.851, see Fig 6b for an example).

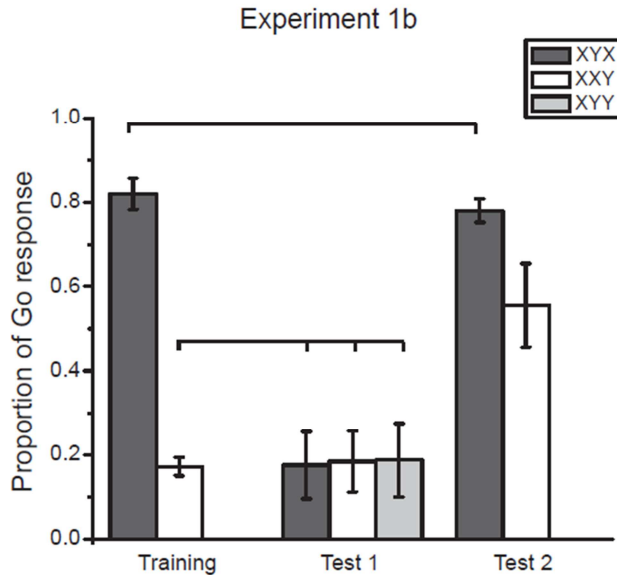


Fig. 5 Mean ($\pm$ SEM, $n = 7$) proportion of Go responses to training and test stimuli. Stimuli in Test 1 consisted of novel elements not present in training. Stimuli in Test 2 consisted of familiar elements from training but with novel combinations. Lines above the bars indicate homogenous responses (no significant differences, X^2 critical value of 12.59) to different stimuli sets.

Discussion

As in Experiment 1a, the zebra finches were able to simultaneously discriminate a large set of different stimuli. Nevertheless, here also the birds did not generalize the discrimination to novel element types not covered in the training sets. It indicates a lack of abstract rule learning, similar to the results in Experiment 1a. However, results of the second test showed that, when novel stimuli consisted of re-arranging familiar elements from the training sets, some birds were able to discriminate between different structures. While the equal response to the XXY training and test stimuli as shown by bird 953 (Fig. 6a) might be because they share the bigrams present in the first two elements, this does not explain the high response that this and the other individuals showed towards the test stimuli with the XYX structure, which do not share bigrams in identical positions to the XXY training stimuli. Note that in fact the XYX test stimuli in this experiment share the final bigrams with the XXY training stimuli (e.g. AB, BC) and the other way around. If the birds were classifying the test stimuli according to the final bigrams, they would have classified XYX test stimuli as XXY and the other way around. However, those birds that did discriminate in the test classified them according to their XYX or XXY structure, hence opposite to the position of the final bigrams. Although these results may indicate that the birds detected, and hence abstracted, the XYX structure when familiar items were used, an alternative, simpler explanation that can not yet be excluded relates to the initial bigrams of the test stimuli. The birds may have been using a rule like: if it starts with AA, BB, CC or DD treat the stimulus as a NoGo stimulus; if these specific bigrams are not present, treat stimuli as a Go stimulus. Attention to repetitions was also demonstrated in an earlier study on zebra finches

using similar stimuli (van Heijningen, Chen et al., 2013), as well as in a recent study on Bengalese finches (Seki, Suzuki et al., 2013). Such a rule would be an example of a stimulus bound generalization (see general discussion).

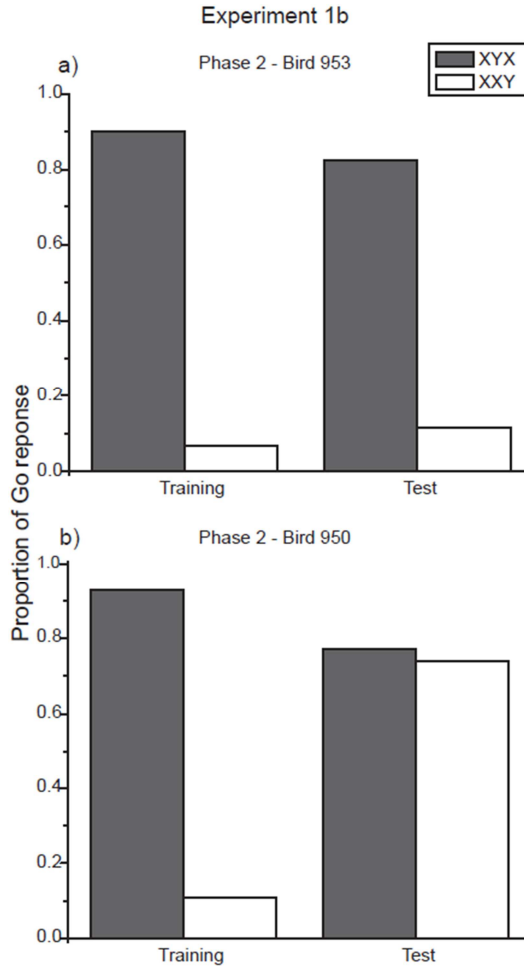


Fig. 6 Individual differences in Go responses to the training and test stimuli in phase 2 of Experiment 1b. a) Bird 953 showed clear pattern of discriminating both training and test stimuli. b) Bird 950 showed no significant discrimination among the test stimuli although it was able to discriminate the training stimuli. Test stimuli were composed with familiar elements known from the training stimuli, but in novel combinations.

Rule learning in humans - Experiment 2a

Experiment 1 demonstrated that zebra finches discriminate the various training sets by concentrating on the presence and position of particular elements, rather than generalizing on the *XYX* or *XXY* patterns. This is unlike the pattern shown by human infants (Marcus, Vijayan et al., 1999). While a limitation on rule learning abilities in birds might be a reason for this difference, another one might be the methodology. Infants were tested using a habituation paradigm, while the current experiment used a Go-NoGo discrimination task. To examine this

issue, Experiment 2 (a and b) subjected adult human participants to the same training and tests as used for the zebra finch Experiments 1a and 1b respectively, except that a categorizing setup was used instead of the Go/No-go.

Experiment 2a was a shortened version of Experiment 1a. It started off with a training and test using the same stimuli as in phase 1of Experiment 1a, but as a pilot experiment showed that humans already discriminated the stimulus sets used in test phase 1 we decided to present all test stimuli immediately after an individual finished training phase 1, without any further training.

Material and Method

Subjects and apparatus

Twenty six native speakers of Dutch (15 males and 11 females) with mean age of 22.12 ± 0.51 SEM (range 19 to 30) were asked to participate. None of them had experience with this kind of experiment. The experiment took place in a quiet area, where participants sat in front of a laptop with a headphone playing sounds at a standardized volume of 40 dB. A custom-designed program (in Processing 1.5.1) controlled training and testing of the participants. At the start of the experiment, the computer screen showed an instruction asking the participants to click on the mouse to trigger a sound playback. After the sound playback a red and a green box were shown. The participants were instructed to categorize the sounds that they heard by clicking one of the boxes, but received no further instructions. They had to discover that a click on the green box after hearing a *XYX* stimulus or on the red box after hearing a *XXY* stimulus resulted in being ‘rewarded’ with a picture of a smiling face. If they clicked on the green box after hearing an *XXY* stimulus or on the red box after hearing an *XYX* stimulus, they were ‘punished’ with a picture of a sad face. Participants had to respond within 5s, otherwise a neutral face was presented and the answer was considered incorrect.

Table 3 Training and test stimuli of Experiment 2a

Training stimuli		Test stimuli (non-reinforced)			
Go (<i>XYX</i>)	NoGo (<i>XXY</i>)	control	<i>XYX</i>	<i>XXY</i>	<i>XYX</i>
		1	BAB	BBA	
			BCB	BBC	
			BDB	BBD	
ABA	AAB	ABA			
ACA	AAC	AAB			
ADA	AAD		EFE	EEF	FEE
		2	FEF	FFE	EFF

Letters indicate different song elements. See legend of Table 1 for further explanation.

Stimuli

The human participants were tested with the same zebra finch song elements and element sequences as used for the zebra finches. However, we increased the interval between the elements to 100 instead of 40 ms, as birds are capable of more precise temporal resolution for complex sounds than human (Dooling, Leek et al., 2002) and our pilot experiment indicated that with the 40 ms interval humans had problems with detecting that the stimuli consisted of three

elements. Table 3 shows the stimulus compositions used in this experiment. As in the zebra finch experiment, we shuffled the element assignment for the participants to avoid coincidental effects caused by particular elements or element combinations.

Procedure

Participants were first trained to categorize the training stimuli. In a first training session they would hear 24 sound sequences (3 *XYX* and 3 *XXY* stimuli each played 4 times). If their responses to the stimuli reached criterion (correct categorization of the two structures each higher than 75%), they would go over to the test phase. If they did not reach criterion performance they would hear another series of 24 sequences. After reaching the training criterion, the participants would see an instruction (in Dutch), saying: ‘Het volgende deel van de test gaat nu beginnen. Je krijgt nu geen feedback’ (The next phase of the test will now start. You will not receive feedback). In the two test phases (phase 2 started right after phase 1), each stimulus was played two times, including two control stimuli that were identical to the training stimuli.

Performance evaluation

In this experiment, participants categorized stimuli according to their structure. Therefore, in order to compare the results directly with the zebra finch experiments, the correct responses to *XYX* stimuli were treated as Go responses to Go stimuli and the incorrect responses to *XXY* stimuli were treated as Go responses to Nogo stimuli. The performance in the final training sessions was compared with that of the tests. In this experiment, the total number of test stimuli per individual participant was relatively low, so we analyzed the data at group level. STP based on G-tests was applied to detect how the participants categorized different stimuli.

Results

Twenty four participants finished the training and the test; two participants quit after being unable to reach the training criterion. The participants needed on average 74 ± 15.95 SEM ($n = 24$) trials to reach criterion.

Fig. 7 shows the results of the test. The participants discriminated the *XYX* and *XXY* structures quite well during the final training session. Although they showed some decrease in discrimination with the test stimuli, the responses to the *XYX* structure and to the *XXY* structure were significantly different for both tests using familiar elements (BAB, etc. $G = 60.86$, $df = 1$, X^2 critical value = 12.59) and unfamiliar elements (EFE, etc. $G = 34.82$, $df = 1$, X^2 critical value = 12.59). However, the responses to the *XXY* and *XYX* structures showed no significant difference ($G = 0.12$, $df = 1$, X^2 critical value = 12.59).

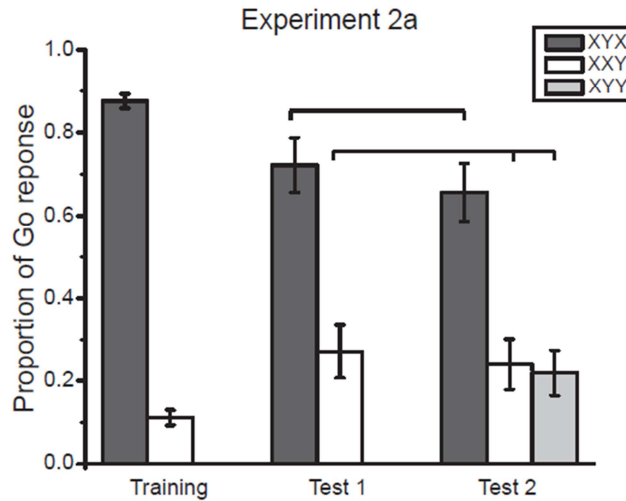


Fig 7 Mean ($\pm$ SEM, $n = 24$) proportion of Go responses in the training (last section) and tests in Experiment 2a. Lines above the bars indicate homogenous responses (no significant differences, X^2 critical value of 12.59) to different stimuli sets.

Discussion

Unlike the zebra finches, the human participants generalized the discrimination to the test stimuli, no matter whether these were composed of familiar elements that they heard during training or of novel elements that they had never encountered before. So, even with a very limited training set, humans already generalized the underlying pattern, rather than using the individual sounds to discriminate the stimuli. In the second test, the responses to XXY and XYY structures were not significantly different, indicating that presence, but not position, of the repeated elements might have been used for the discrimination. Indeed, when the participants were asked what pattern they had detected in the stimuli, most mentioned that they paid attention to the presence or absence of repeated elements to discriminate the stimuli.

Rule learning in humans - Experiment 2b

This experiment used the same training and test sequences as used in Experiment 1b.

Material and Method

Subjects and apparatus

Another twenty six native Dutch speakers (15 males and 11 females), with a mean age of 22.15 ± 0.5 SEM (range 19 to 30) and no experience with this kind of experiment, participated in this experiment. The training and testing conditions were identical to Experiment 2a.

Stimuli

The stimuli for this part of experiment were composed in the same way as those in Experiment 1b, but, as in Experiment 2a, using a 100 ms instead of a 40 ms pause in between the elements. There was one training phase with 6 pairs of stimuli and two test phases with 3 and 4 pairs of stimuli respectively, identical to the layout of Experiment 1b (Table 2). Element assignments for the participants were shuffled.

Procedure

The procedure was identical to Experiment 2a except that different training and test stimuli were used (see Table 2).

Performance evaluation

The same statistical methods as for Experiment 2a were used in this experiment.

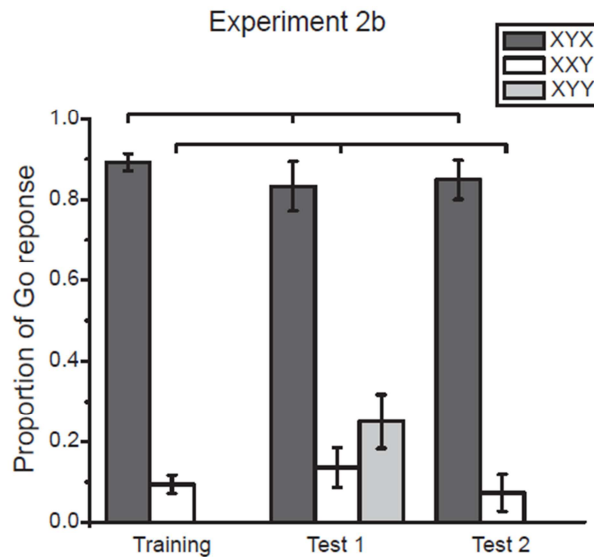


Fig. 8 Mean ($\pm$ SEM, $n = 24$) proportion of Go responses in the training (last section) and tests in Experiment 2b. Lines above the bars indicate homogenous responses (no significant differences, X^2 critical value of 12.59) to different stimuli sets.

Results

Two out of 26 participants did not finish the experiment due to the failure during training. The other 24 participants reached criterion performance after on average 89 ± 11.42 (Mean $\pm$ SEM) trials. As shown in Fig 8, participants showed no significant differences in responding to the XYX stimuli in the last section of training and XYX stimuli presented in the both tests ($G = 3.10$, $df = 2$, X^2 critical value = 12.59). Also the responses to XXY stimuli did not differ between training and tests ($G = 2.80$, $df = 2$, X^2 critical value = 12.59). However, the

responses to the XYY stimuli belonged to neither this XYX group ($G = 154.35$, $df = 3$, X^2 critical value = 12.59) nor to the XXY group ($G = 18.84$, $df = 3$, X^2 critical value = 12.59).

Discussion

Again, the human participants discriminated novel XYX from novel XXY strings. Compared with the performance of Experiment 2a, participants had higher accuracies in categorizing the stimuli according to their structure. In addition, participants categorized the XYY stimuli different from the other ones, although they categorized XYY stimuli more like XXY stimuli than like XYX stimuli. This outcome shows that the participants not only used the element repetitions as criterion to categorize the stimuli, but also attended to the position of the repeated elements. Again, this was confirmed by the comments of the participants after the experiment.

General discussion

Our results demonstrate that birds and humans show different levels of generalization. The results of Experiment 1 demonstrate that zebra finches are able to categorize element strings according to their structure, i.e. they pay attention to the sequence and position of elements, confirming results obtained in earlier studies in this species (van Heijningen, Chen et al., 2013; van Heijningen, de Visser et al., 2009), as well as in Bengalese finches (Seki, Suzuki et al., 2013). This ability to discriminate between XYX and XXY strings is not based on rote memorization of individual strings. If so, they would treat all test stimuli as equally novel. Rather, the birds showed a generalization in which they used familiar and indicative elements (the identity -A, B, C or D- of the final element in the XYX stimuli in the phases 1-4 of Experiment 1a) or element bigrams (AA, BB, CC, DD in Experiment 1b) to classify novel stimuli. This generalization is stimulus bound in the sense that the birds can track a regularity in a surface feature, and only generalize to strings with the indicative elements in the same position as in the training strings and not to strings composed from novel elements. These results show a striking similarity to those obtained in an experiment by Gerken (2006) on 9-months old human infants. Infants exposed to a XYX or XXY strings of consonant-vowel (CV) items in which they could use the presence or absence of a particular CV item in a specific position (i.e. the Y-item was the same for all training stimuli, but the X varied) generalized to novel strings that had the same Y-item in the same position as in the training stimuli. They did not generalize to novel items having the same abstract structure, but in which both X and Y were novel. It was only when the infants were trained with several different strings in which both X and Y differed between the strings (e.g. ABA, CDC, etc.) that they were able to generalize to novel items. A later study (Gerken, 2010) showed that even a limited exposure to such strings was sufficient to induce the generalization. It is here that the results of the infants and zebra finches diverge: zebra finches trained with a set of stimuli that share their abstract structure only (Experiment 1b) did not generalize to novel stimuli and hence did not detect the underlying pattern. The birds in our Experiment 1a and the infants in the experiment described by Gerken (2006) were exposed to input that can be described at at least two levels. One is the more abstract level based on the string structure (XYX or XXY), another is a simpler one that is restricted to learning that a particular element has a particular position in the strings. Given that the zebra finches in our

experiments failed on learning the more abstract structure even though the training strings in Experiment 1b are suitable for learning such a pattern, suggests that we have, at least in this experiment, reached a limit on the abilities for generalization in zebra finches. Whether this demonstrates an absolute barrier or whether such generalization might be induced by training with a larger, more varied training set remains to be investigated. The methodology might also restrict the learning of a more complex generalization. As in most comparison studies between humans and other animals (e.g. Stobbe et al. 2012), the human subjects in our experiment required far fewer trials than zebra finches. This might have affected what type of regularities have been learned Peña, Bonatti et al. (2002). In the Go/Nogo paradigm, the animals require extensive training, which might focus attention towards the familiar elements or bias rote memorization.

To demonstrate the acquisition of abstract rules, successful categorization should not depend on the physical similarity of items but should also be present when the exemplars bear no physical similarity to each other (Lea, 1984; Pearce, 2008). Human adults in our experiment and infants in the study by Marcus, Vijayan et al. (1999) were successful in categorizing novel strings into different groups according to their structures. This kind of task may require knowledge of a second-order relationship, in which the subjects need to detect the relationship between two other relationships (Thompson & Oden, 2000). Given an exemplar like ABA and a test stimulus like EFE, learners needed to know the relationships between individual elements within each sequence and the relationship between the two sequences to enable correct generalizing. Earlier studies showed that starlings (Comins & Gentner, 2013; Gentner, Fenn et al., 2006), zebra finches (van Heijningen, de Visser et al., 2009) and Bengalese finches (Seki, Suzuki et al., 2013) are all able to generalize a string structure to a novel string in which the individual items are replaced by novel ones from the same phonetic category. Of these species, zebra finches are the only ones that have been tested with the novel strings consisting of items belonging to novel categories and they fail in such a generalization. This failure might indicate that this ability, which is fundamental to syntactic abilities, is lacking or only marginally present, as is concluded from other studies on various non-human animals (Oden, Thompson et al., 1990; Thompson & Oden, 2000). However, a few individual birds in the current study succeeded in discriminating between stimulus strings that consisted of familiar element types but novel combinations, similar to what has been found in rhesus monkeys (Hauser & Glynn, 2009). While it cannot be excluded that this generalization was based on a rule like: ‘if the string contains AA, BB, CC or DD bigrams treat it as a NoGo-stimulus, if it does not, treat as a Go-stimulus’, it might also indicate that some abstract learning has occurred, but linked to the presence of familiar items (song elements).

Against the above background, the results of the human adults are of particular interest. In contrast to both birds and infants they readily generalized to novel strings even after training with a single set of XYX and XXY stimuli, although the stimulus set allowed also for learning the same regularity as the zebra finches (the final element of the XYX strings). The difference between adults and infants might be related to their experiences: human adults have been trained to master rule abstraction over their life. But even though humans thus show a broader generalization than the birds do, they still base their discrimination on applying a relatively

simple rule: presence or absence of repeated elements (XX) rather than on learning an XYX rule. Such a bias for learning about element repetitions is also shown in human infants (Gervain, Macagno et al., 2008; Kovács & Mehler, 2009). This might reflect a bias to use so-called perceptual or memory primitives (POMPs), as suggested by Endress, Nespor et al. (2009). Only after the more extensive training with more exemplars, as in Experiment 2b, did human adults also recognize that the structure is XXY and not XYX and hence applied a more extensive rule. Interestingly, den Bos and Poletiek (2010) using an artificial grammar task in which human adults could use either a grammar cue derived from example strings or a more local cue (a particular item in a fixed position in strings), also found their participants to focus on the local cue over the grammar one.

In the current study we found that zebra finches shared the same learning mode with human infants when they could use a stimulus bound generalization. However, while human adults succeeded in abstracting structural rules from the stimuli, abstracting in zebra finches seems constrained to the identity and position of specific diagnostic stimulus items.

To conclude, simple rule abstraction may be shared between humans and birds, but it remains to be investigated at which level of abstraction rule learning of non-human animals, whether mammals or birds, diverges from that of humans. Without implying that the abilities as observed in zebra finches are phylogenetically widespread or representative of other non-human species, the results do suggest a scenario in which the human ability to derive and apply a more general or complex underlying rule from experience with specific exemplars is something that might have arisen, both in ontogeny and during evolution, in a gradual progression from identifying individual stimuli, via stimulus bound generalization and recognizing patterns like repetitions towards more abstract patterns.

References

- Altmann, G. T. M., & Dienes, Z. (1999). Rule Learning by Seven-Month-Old Infants and Neural Networks. *Science*, 284(5416), 875. doi: 10.1126/science.284.5416.875a
- Beckers, G. J., Bolhuis, J. J., Okanoya, K., & Berwick, R. C. (2012). Birdsong neurolinguistics: songbird context-free grammar claim is premature. *Neuroreport*, 23(3), 139-145. doi: 10.1097/WNR.0b013e32834f1765
- Berwick, R. C., Okanoya, K., Beckers, G. J., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in cognitive sciences*, 15(3), 113-121. doi: 10.1016/j.tics.2011.01.002
- Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, 11(11), 747-759. doi: 10.1038/Nrn2931
- Comins, J. A., & Gentner, T. (2014). Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. *Animal cognition*, 1-8. doi: 10.1007/s10071-014-0732-5
- Comins, J. A., & Gentner, T. Q. (2013). Perceptual categories enable pattern generalization in songbirds. *Cognition*, 128(2), 113-118. doi: 10.1016/j.cognition.2013.03.014
- Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive science*, 31(4), 697-704. doi: 10.1080/15326900701399947
- Corballis, M. C. (2009). Do rats learn rules? *Animal Behaviour*, 78(4), E1-E2. doi: 10.1016/j.anbehav.2009.05.001
- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5), 1119-1130. doi: 10.1037/0278-7393.30.5.1119
- 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. doi: 10.1016/j.cognition.2012.09.015
- den Bos, E., & Poletiek, F. (2010). Structural selection in implicit learning of artificial grammars. *Psychological Research PRPF*, 74(2), 138-151. doi: 10.1007/s00426-009-0227-1
- Dooling, R. J., Leek, M. R., Gleich, O., & Dent, M. L. (2002). Auditory temporal resolution in birds: discrimination of harmonic complexes. *The Journal of the Acoustical Society of America*, 112(2), 748-759.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. doi: 10.1146/annurev.neuro.22.1.567
- Eimas, P. D. (1999). Do infants learn grammar with algebra or statistics? *Science*, 284(5413), 435-436; author reply 436-437.
- Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica*, 135(2), 182-190. doi: 10.1016/j.actpsy.2010.06.005
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal cognition*, 13(3), 483-495. doi: 10.1007/s10071-009-0299-8
- Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, 105(3), 577-614. doi: 10.1016/j.cognition.2006.12.014
- Endress, A. D., Nespore, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in cognitive sciences*, 13(8), 348-353. doi: 10.1016/j.tics.2009.05.005
- Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1933-1955. doi: 10.1098/rstb.2012.0103
- Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303(5656), 377-380.
- Folia, V., Udden, J., de Vries, M., Forkstam, C., & Petersson, K. M. (2010). Artificial Language Learning in Adults and Children. *Language Learning*, 60, 188-220. doi: 10.1111/j.1467-9922.2010.00606.x
- Frank, M. C., Slemmer, J. A., Marcus, G. F., & Johnson, S. P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Developmental science*, 12(4), 504-509. doi: 10.1111/j.1467-7687.2008.00794.x
- Gabriel, K. R. (1966). Simultaneous Test Procedures for Multiple Comparisons on Categorical Data. *Journal of the American Statistical Association*, 61(316), 1081-&. doi: 10.2307/2283202
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic bulletin & review*, 16(3), 486-490. doi: 10.3758/PBR.16.3.486
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088), 1204-1207. doi: 10.1038/nature04675
- Gerken, L. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition*, 98(3), B67-B74. doi: 10.1016/j.cognition.2005.03.003
- Gerken, L. (2010). Infants use rational decision criteria for choosing among models of their input. *Cognition*, 115(2), 362-366. doi: 10.1016/j.cognition.2010.01.006

- Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, 105(37), 14222-14227. doi: 10.1073/pnas.0806530105
- Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature*, 410(6831), 930-933. doi: 10.1038/35073582
- Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in cognitive sciences*, 4(5), 178-186.
- 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(5598), 1569-1579. doi: 10.1126/science.298.5598.1569
- Hauser, M. D., & Glynn, D. (2009). Can free-ranging rhesus monkeys (*Macaca mulatta*) extract artificially created rules comprised of natural vocalizations? *Journal of Comparative Psychology*, 123(2), 161-167. doi: 10.1037/a0015584
- 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(3), B53-64.
- Hauser, M. D., Weiss, D., & Marcus, G. (2002). RETRACTED: Rule learning by cotton-top tamarins. *Cognition*, 86(1), B15-B22.
- Herbranson, W. T., & Shimp, C. P. (2003). "Artificial grammar learning" in pigeons: a preliminary analysis. *Learning & behavior*, 31(1), 98-106.
- Herbranson, W. T., & Shimp, C. P. (2008). Artificial grammar learning in pigeons. *Learning & behavior*, 36(2), 116-137.
- Katz, J. S., & Wright, A. A. (2006). Same/different abstract-concept learning by pigeons. *Journal of experimental psychology. Animal behavior processes*, 32(1), 80-86. doi: 10.1037/0097-7403.32.1.80
- Kovács, Á. M., & Mehler, J. (2009). Flexible Learning of Multiple Speech Structures in Bilingual Infants. *Science*, 325(5940), 611-612. doi: 10.1126/science.1173947
- Lachlan, R. F., Verhagen, L., Peters, S., & ten Cate, C. (2010). Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taenopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *Journal of Comparative Psychology*, 124(1), 92-108. doi: 10.1037/a0016996
- Lea, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblat, T. G. Bever & H. S. Terrace (Eds.), *Animal cognition* (pp. 263-276). Hillsdale, NJ: Erlbaum: Psychology Press.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77-80.
- Mondragon, E., Murphy, R. A., & Murphy, V. A. (2009). Rats do learn XYX rules. *Animal Behaviour*, 78(4), E3-E4.
- Murphy, R. A., Mondragon, E., & Murphy, V. A. (2008). Rule learning by rats. *Science*, 319(5871), 1849-1851. doi: 10.1126/science.1151564
- Negishi, M. (1999). Do infants learn grammar with algebra or statistics? *Science*, 284(5413), 435; author reply 436-437.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, 49(2), 85-117. doi: 10.1016/j.cogpsych.2003.12.002
- Oden, D. L., Thompson, R. K. R., & Premack, D. (1990). Infant Chimpanzees Spontaneously Perceive Both Concrete and Abstract Same Different Relations. *Child Development*, 61(3), 621-631. doi: 10.1111/j.1467-8624.1990.tb02807.x
- 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. *Animal cognition*, 15(2), 155-161. doi: 10.1007/s10071-011-0441-2
- Pearce, J. M. (2008). *Animal Learning and Cognition*. New York: Psychology Press.
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, 298(5593), 604-607.
- Pinker, S. (1991). Rules of Language. *Science*, 253(5019), 530-535. doi: 10.1126/science.1857983
- Reber, A. S. (1967). Implicit Learning of Artificial Grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 855-863. doi: 10.1016/S0022-5371(67)80149-X
- Reber, A. S. (1990). On the Primacy of the Implicit - Comment. *Journal of Experimental Psychology: General*, 119(3), 340-342. doi: 10.1037//0096-3445.119.3.340
- 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(1), 27-52.
- 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(3), 669-680. doi: 10.1016/j.cognition.2006.11.004
- Seidenberg, M. S. (1999). Do infants learn grammar with algebra or statistics? *Science*, 284(5413), 434-435; author reply 436-437.

- Seki, Y., Suzuki, K., Osawa, A. M., & Okanoya, K. (2013). Songbirds and humans apply different strategies in a sound sequence discrimination task. *Frontiers in psychology*, 4, 447. doi: 10.3389/fpsyg.2013.00447
- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: the principles and practice of statistics in biological research*. New York: Freeman and Co.
- 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*). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1995-2006. doi: 10.1098/rstb.2012.0096
- ten Cate, C., Lachlan, R., & Zuidema, W. (2013). Analyzing the structure of bird vocalizations and language: finding common ground. In J. J. Bolhuis & M. Everaerts (Eds.), *Birdsong, speech, and language* (pp. 243-260). Cambridge: The MIT press.
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1984-1994. doi: 10.1098/rstb.2012.0055
- Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive science*, 24(3), 363-396. doi: 10.1207/s15516709cog2403_2
- Toro, J. M., & Trobalon, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & psychophysics*, 67(5), 867-875.
- 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(2), 165-175. doi: 10.1007/s10071-012-0559-x
- 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. *Proceedings of the National Academy of Sciences*, 106(48), 20538-20543. doi: 10.1073/pnas.0908113106
- Wright, A. A., Rivera, J. J., Katz, J. S., & Bachevalier, J. (2003). Abstract-concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of experimental psychology. Animal behavior processes*, 29(3), 184-198.
- Yamazaki, Y., Suzuki, K., Inada, M., Iriki, A., & Okanoya, K. (2012). Sequential learning and rule abstraction in Bengalese finches. *Animal cognition*, 15(3), 369-377. doi: 10.1007/s10071-011-0462-x

Chapter 5^{*}

Bridging the gap: learning of non-adjacent dependencies in a songbird

^{*} This chapter is based on: Jiani Chen & Carel ten Cate (manuscript).

Abstract

Detecting dependencies between linguistic items is crucial for language learning. Compared to adjacent dependency acquisitions, learning non-adjacent dependencies is more challenging and might require more complex cognitive skills, as non-adjacent dependencies require detection of a relation between separated items irrespective of the number and the nature of the intervening events. So far, only two non-human primates were shown to learn non-adjacent dependencies, both between items present at the edge positions of a sound sequence. However, dependencies need not necessarily concern the edges of a sequence. In the current study, zebra finches were trained and tested to examine their ability to learn a non-adjacent dependency over a variable number and over different types of intervening items. A separate experiment addressed whether a zebra finch could learn non-adjacent dependencies between items present at arbitrary positions of a sequence. The results show that the birds cannot only learn non-adjacent dependencies at an arbitrary distance and over novel intervening items but can also learn non-adjacent dependencies at arbitrary positions of a sequence. Our findings show an ability for detecting a grammatical pattern thus far not demonstrated for a non-human animal.

Introduction

Language is hierarchically organized system and in its complexity unique to humans. Although there are many differences between human language and non-human animal communication systems (Berwick, Okanoya et al., 2011; Hauser, Chomsky et al., 2002), recent research has revealed that some cognitive components, which are critical to the linguistic system, may be shared with non-human animals (reviewed in Bolhuis & Everaert, 2013; Bolhuis, Okanoya et al., 2010). This may indicate domain general cognitive properties that have been at the basis of the language faculty. Studying the learning abilities that are shared between humans and non-human animals not only reveals what other animals can do but can thus also provide insights in the evolution of language.

Combinatorial syntax has been considered to lie at the center of the distinctive and creative nature of languages (Berwick, Okanoya et al., 2011). One indispensable skill for syntactical learning is detecting the relations between words and morphemes in a sentence. For example, in English, speakers need to consider dependencies involving number agreement and tense. Much work has focused on adjacent dependencies and several studies showed that learning such dependencies are rapidly acquired and domain general. Human infants exposed to artificial language for two minutes could detect adjacent dependencies between syllables (Saffran, Aslin et al., 1996). Subsequently, this ability was also found in nonhuman primates (Hauser, Newport et al., 2001).

Sensitivity to more sophisticated dependencies between words, like dependencies that are not adjacent in the real-time linear sequence of a sentence (e.g. **the bird** the cat chases **flies away**), is necessary for learning the hierarchical structure of language (Lany & Gómez, 2008). Compared to adjacent dependencies, non-adjacent ones are more challenging for learners, since they require learners to detect the relationships over intervening material. Such dependencies are not easy to acquire. Human infants can learn adjacent dependencies at a young age (8-month old or even younger, Saffran, Aslin et al., 1996; Thiessen & Saffran, 2003) but fail to track non-adjacent ones before 15 months of age (Gómez & Maye, 2005; Santelmann & Jusczyk, 1998). Even adults have shown restricted success in acquiring non-adjacent dependencies in artificial languages (Gómez, 2002; Newport & Aslin, 2004). Opinions vary as to whether the detection of non-adjacent dependencies relies on abstract, algebraic like computations or whether more simple associative mechanisms may explain the phenomenon (Perruchet, Tyler et al., 2004) and whether the phenomenon is limited to language learning or of a more general nature. This latter issue makes it of interest to address the question whether the ability to detect non-adjacent dependencies is present in non-human animals. There is evidence suggesting that some non-human primates can become aware of non-adjacent dependencies between sounds. In a study of Newport, Hauser et al. (2004), tamarin monkeys were able to detect certain types of non-adjacent dependencies between the first and third item in three-item strings. The tamarins differentiated different non-adjacent dependencies when the strings were constructed from three CV (consonant-vowel) syllables or when the items were arranged in VCV strings, but not to strings in CVC structure. Another species, the squirrel monkey, was able to discriminate between tone sequences containing a non-adjacent dependency spanning 1-5 items and those

lacking it (Ravignani, Sonnweber et al., 2013). This indicates that some cognitive ability to detect non-adjacent dependencies might be present in other animals.

Experiments demonstrating some evidence for the ability to detect non-adjacent dependencies were all in primates up to now (one study in rats failed to find such learning, Toro & Trobalon, 2005) and the non-adjacent dependencies in non-human animal studies all occurred between items present at the edges of a sequence of sound units. Edge positions of a sequence are considered to be salient and easier to detect (Endress, Carden et al., 2010; Endress, Nespor et al., 2009; Henson, 1998), therefore this may reduce the difficulty of learning non-adjacent dependencies. In human language, dependencies do not necessarily occur at the edges. This raises the questions whether learning of non-adjacent dependency requires higher cognitive skills that only occur in primates and whether nonhuman animals can learn non-adjacent dependencies between items present at arbitrary positions in a sequence. In the current study, we explore these questions by using a songbird species, the zebra finch. Songbirds, such as zebra finches, share striking neural, behavioral and genetic similarities in auditory-vocal learning with humans (Bolhuis, Okanoya et al., 2010; Doupe & Kuhl, 1999; Fisher & Scharff, 2009; Jarvis, Güntürkün et al., 2005). Their songs are also hierarchically organized, but there is no evidence yet that they contain non-adjacent dependencies. However, this need not imply that birds cannot detect such dependencies, as this may be a more general cognitive ability, not strictly linked to vocal production.

We trained and tested zebra finches with strings derived from an artificial grammar that can generate strings containing different non-adjacent dependencies. In experiment 1, elements that depended on each other always occurred at the edges of sequences but the number of intervening elements in the middle varied from 1 to 3. We also tested whether the ability to detect the dependencies remained when the nature of the intervening items changed. One step further, in experiment 2, we tested the ability of this songbird species to detect dependencies between elements that occurred at arbitrary positions in the sequences. For instance, a sequence with a dependency between item A and B could be arranged as AIIBI or IAIBI (I is an intervening item). This has so far not been examined in a non-human species. In the study of Ravignani, Sonnweber et al. (2013), on squirrel monkeys, non-adjacent dependencies were constructed by using one dependent pair of items. This raises the question like whether the animals really detected the non-adjacent dependency or just paid attention to lower level cues, such as the number of transitions within a string. In the current study, we used two pairs of dependent items (item A paired with item B and item C paired with item D) to construct non-adjacent dependent stimuli. Correct classification of the stimuli could not be done by attending to the number of element transitions, nor by attending to the nature of the item at just to one of the edges. Our results reveal the ability of zebra finches to learn non-adjacent dependencies, an ability previously unknown to be present in songbirds. We also provide evidence suggesting that zebra finches can even generalize dependencies that occur at arbitrary positions in sequences.

Materials and methods

Subjects and apparatus

Subjects were 6 zebra finches from the Leiden University breeding colony, 3 females and 3 males. They were all naïve to any experimental setup. Five birds participated in Experiment 1 and one other bird, which was relatively quick in mastering the first non-adjacent task, was tested in a more complex design (Experiment 2). During the experiment, the birds were individually housed in operant conditioning cages (70×30×45 cm), each of which was placed in a sound attenuated chamber with daylight spectrum light on a 13.5 L: 10.5 D schedule. The cages were made from wire mesh except for the floor and a plywood back wall. The floor was covered with sand. The back wall contained a food hatch and two red pecking sensors. Each sensor contained a red LED that indicated the activated stage of the sensor. Five perches were mounted on the back wall to enable hopping behavior. The sensors and the food hatch could be reached easily from the middle perch. Acoustic stimuli were delivered from a loudspeaker located above the cage and were calibrated to a pressure level of 70 dB (see van Heijningen, de Visser et al., 2009 for further details).

Stimuli

Stimuli were constructed from ramped and RMS equalized zebra finch song elements. Seven types of song elements were selected from our zebra finch song database based on optimal discriminability. Natural zebra finch songs consist of a series of song elements and can be composed of different elements or sung in different element orders by different males (Lachlan, Verhagen et al., 2010; Zann & Bamford, 1996). There is no indication that single elements or element combinations carry meaning and hence these elements are suitable for constructing artificial meaningless strings. Moreover, song elements are not as simple as artificial tones and may be perceived better by birds in comparison to human syllables. Thus using them to construct strings of items excludes the exploitation of simpler cues of sounds while facilitating the recognition for the birds.

In the current study, we generated stimulus sequences based on an artificial grammar (Fig. 1). We created 5 items: A, B, C, D and I. In order to gain a reward, the birds had to learn that item A was paired to item B and item C was paired to item D. Item I was defined as an intervening item occurring between (or before or after) these two dependent pairs. Stimuli generated from this grammar (further indicated as ‘grammatical’ stimuli) were used as Go stimuli in the training. The some five items were also used to generate other element sequences that do not follow the grammar, like AIID or ICIBI. These ungrammatical stimuli were used as Nogo stimuli during the training. Note that this structure also prevented the birds from just using the identity of the first or last element to decide whether a sequence was correct or not, as, for instance, a sequence starting with an A could end with a B (correct) but also with a D (incorrect). To eliminate pseudo effects due to an arbitrary characteristic of the sounds, the song element assignments to A, B, C, D and I were shuffled for the subjects, thus, for instance, item A was a different element type for different birds.

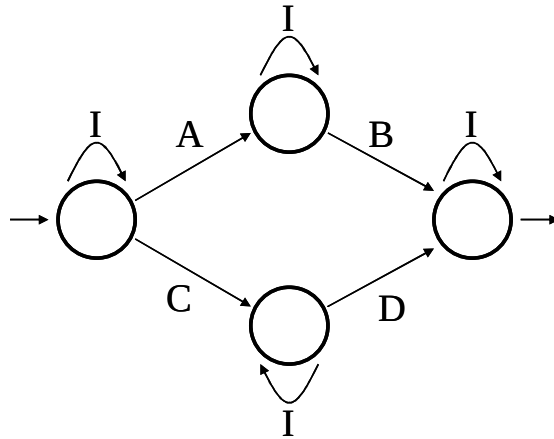


Fig. 1 Diagram of an artificial grammar that generates $I^n A I^n B I^n$ and $I^n C I^n D I^n$ ($n \geq 0$, 1-3 indicate that n is not necessarily the same) sequences that are used in Experiments 1 and 2. Transitions from one state to another (arrow from one circle to another) produce an element of the stimulus sequences.

Procedure

The subjects were trained and tested in a Go/Nogo procedure. They were trained to peck on one sensor to trigger a sound playback and to peck or withhold pecking on the second sensor according to the sound stimulus. If the stimulus was a Go sound, pecking on the second sensor resulted in access to food behind the food hatch for 10s and if the sound was a Nogo sound, pecking on the second sensor would lead to a mild penalty (light switched off for 15s). To familiarize the birds with the Go/Nogo task, they were first trained to discriminate a natural song from a pure tone.

Training with artificial stimuli started when the birds' responses reached the training criterion ($>75\%$ response to Go stimuli and $<25\%$ response to Nogo stimuli). If birds had a persistent low response ratio to Go stimuli or high response to Nogo stimuli, a more lenient criterion was applied (discrimination ratio $>75\%$, calculated as the percentage response to Go stimuli divided by the sum of the percentage response to Go stimuli and the percentage of response to the Nogo stimuli). Stimuli were presented in a random order. Throughout the entire experiment, birds received two types of stimuli: training stimuli and test stimuli. Responding to test stimuli was not reinforced in order to avoid additional learning of these stimuli. To prevent extinction of the pecking behavior during the test phase, training stimuli were presented and reinforced on 80% of the trials and test stimuli were presented on 20% of the trials. When each test stimulus reached 40 trials, tests were stopped and the birds were switched to the next phase.

Experiment 1

In the first experiment, birds were trained and tested with the mutually dependent elements at the edges, e.g. one of the paired items occurred at the beginning of a sequence while the other one occurred at the end of the sequence (Fig. 2, appendix Table 1). The entire experiment consisted of five phases. In the 1st phase, training was 'starting small'. Instead of

training birds with non-adjacent dependencies, adjacent dependencies were used. Two adjacent dependencies (AB, CD) generated from the grammar (Fig. 1) were presented as Go stimuli. Another 6 Ungrammatical stimuli served as Nogo stimuli. Starting from adjacent dependencies is based on the ‘starting small’ theory (Elman, 1993), which entails that limited or reduced input may confer a computational advantage for learning dependencies. Several empirical studies support this claim (Cochran, McDonald et al., 1999; Lai, 2013; Lany & Gómez, 2008; Lany, Gómez et al., 2007).

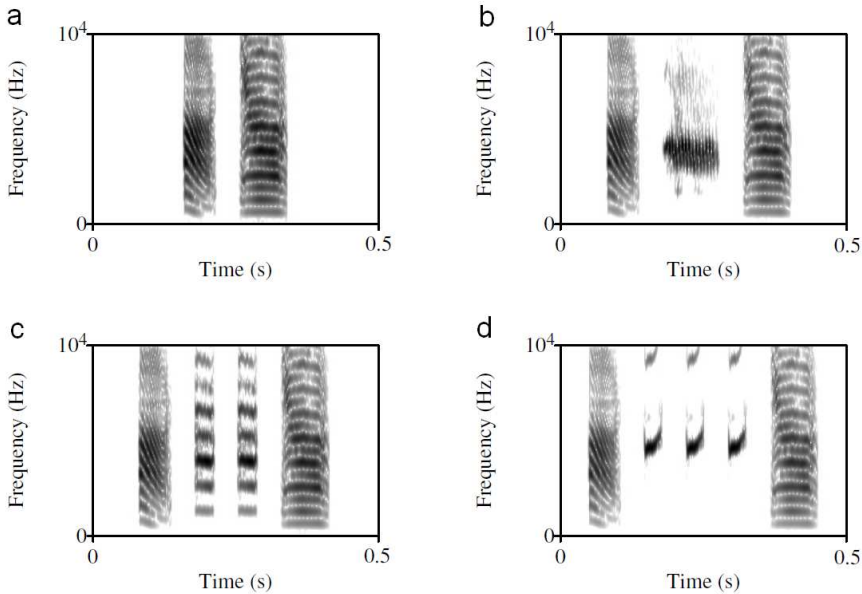


Fig. 2 Spectrograms of four examples of constructed song stimuli for bird 143 in Experiment 1. a) elements arranged in AB order. b) elements in AI₁B order. c) elements in AI₂I₂B order. d) elements in AI₃I₃I₃B order.

After training with adjacent dependencies, birds were tested with stimuli with one intervening element in between the two paired items to check whether they could transfer the dependencies. For instant, after training with stimuli AB and AD etc., in the 1st test birds received test stimuli that also consisted of grammatical sequences (such as AIB) and ungrammatical sequences (such as AID). After completing this test, the birds were next trained with the stimuli from the 1st test during the 2nd training phase. A 2nd test followed again when training reached criterion. The number of intervening items in test 2 was expanded from 1 to 2 and 3, yielding non-adjacent dependencies at varied distances. In addition, another test (Test 2+) was given to check whether the birds could still detect the dependencies when one or both dependent items did not occur at the edges. In the 3rd phase, birds were switched to training with the stimuli used in test 2. In the 3rd test, intervening items of the training stimuli were replaced with a novel element type that the birds had never heard in the previous phases (I₁ elements were replaced with I₂ elements). After this test, a 4th training took place now using the stimuli from the 3rd test as training stimuli. This was followed by a 4th test with intervening items again being

replaced with another novel type of element (I_2 to I_3). Finally, in the 5th phase, the birds were trained with stimuli from the 4th test to examine their learning curves.

In this procedure, training and testing built up gradually towards longer strings (appendix Table 1). Although the distances and the intervening elements between the non-adjacent dependencies varied, the dependencies were always kept the same. If birds could learn such dependencies, they should be able to ignore the intervening elements, no matter how many or of what types these were, and to discriminate grammatical from ungrammatical stimuli. If so, we expected them to transfer the information of the dependencies from one training phase to the next and to maintain the discrimination between grammatical and ungrammatical stimuli when the training switched to a new phase with novel sequences.

Experiment 2

Experiment 2 tested the detection of non-adjacent dependencies at arbitrary positions of a sequence (appendix Table 2). The 1st phase of Experiment 2 was identical to that of Experiment 1. Training was also ‘starting small’ from adjacent dependencies and the first test was with a single I-element inserted. The 2nd training was also identical to that of Experiment 1. However, in the stimuli of test 2, items that formed non-adjacent dependencies occurred at arbitrary positions in a 5 element sequence, such as AIIBI and ICIDI. In these stimuli, the edges did not predict the dependencies. The number of intervening items also varied, resulting in non-adjacent dependencies at variable distances. Twelve grammatical stimuli and sixty ungrammatical stimuli with non-adjacent dependencies could be generated. Five stimuli of each type (grammatical and ungrammatical) were used in this test. A 3rd training followed after test 2. Another 2 grammatical and ungrammatical stimuli were added to the stimuli from the 2nd test and used as training stimuli. After this training, 4 grammatical and 10 ungrammatical stimuli that were structured differently from those used in the training were presented in the 3rd test. Subsequently one more test (test 4) was done, in which the intervening elements of the stimuli were replaced by a novel type (I_1 to I_2). Training 4 started when test 4 was completed, using the stimuli from test 4 as training stimuli. Similarly, in test 5 the intervening elements of stimuli were once again replaced with a novel type (I_2 to I_3). The bird was switched to a final training after test 5, using the stimuli of test 5 for training. We expected results similar to those mentioned for Experiment 1.

Statistics

A logistic regression model was applied to the learning data of every individual bird by using binary data (response/ non-response) as dependent variable and trials as independent variable to construct the learning curves of the grammatical and ungrammatical stimuli. The first 4800 trials were taken into the model analysis, because most birds reached the lower criterion of training (discrimination ratio >75%) when they received 4800 trials of training. For significance testing of the test data, we used a G-test (with Williams correction if the number of responses was < 5) to analyze the difference in responses to the grammatical and ungrammatical stimuli. Within the ungrammatical set, more detailed analyses were conducted to check whether there were differences between responses to different ungrammatical structures. In the ungrammatical

set, stimuli could be categorized basically into three categories. We defined the first category as Co-Occurrence (CO), since they contained both paired items but the transitions between paired items were incorrect, such as BIIA, DIIIC. We named the second one as Position-Correct (PC), of which the edge positions of the stimulus sequences coded correctly but the syntactical dependencies between items were incorrect, such as AIID, CIIIB. The third category contained other stimuli (Other) that did not occur in the first two categories. Simultaneous Testing Procedures (STP) based on a G-test were applied when stimuli categories were more than two (Gabriel, 1966; Sokal & Rohlf, 1995).

Results

Experiment 1. Edge based non-adjacent dependency

Three out of 5 birds completed the experiment. One bird died in the beginning of the experiment and one bird could not discriminate the grammatical and ungrammatical bigrams (AB and CD ect.) of the first training, even after a long training period.

Phase 1: During the 1st training, birds were trained with adjacent dependencies consisting of bigram stimuli. They gradually learned to distinguish grammatical from ungrammatical stimuli when the training progressed (Fig. 3 Training 1, appendix Table 3). After having reached the training criterion, the birds were given Test 1, in which the dependencies between items A and B and items C and D became non-adjacent dependencies by inserting one I-element between the items. No significant differences between responses to grammatical and ungrammatical stimuli were found for bird 150 ($G = 2.26$, $df = 1$, $P = 0.13$) and bird 157 ($G = 0.58$, $df = 1$, $P = 0.45$). However, the bird 143 showed significantly more responses to the grammatical than to the ungrammatical stimulus set ($G = 4.42$, $df = 1$, $P = 0.04$, Fig. 4 Test 1). Responses to different categories within ungrammatical stimuli showed no significant difference (bird 150: $G = 1.23$, $df = 1$, $P = 0.27$; bird 143: $G = 0.38$, $df = 1$, $P = 0.54$; bird 157: $G = 0.81$, $df = 1$, $P = 0.37$).

Phase 2: In the 2nd training, an intervening element was inserted into the bigram stimuli from the 1st training to generate non-adjacent dependencies. Results of the 2nd training showed that all three birds, and in particular bird 143 and 150, seemed to transfer what they learned previously to the new task. They learned to discriminate the grammatical and ungrammatical sets at the beginning of the training (Fig. 3 Training 2, appendix Table 3). Learning was faster in comparison with learning during the 1st training.

After training with non-adjacent dependencies with 1 intervening element, birds were tested with non-adjacent dependencies with 2 and 3 intervening elements. Now two birds responded significantly more to the grammatical than to ungrammatical stimuli (bird 150: $G = 52.64$, $df = 1$, $P < 0.001$; bird 143: $G = 12.98$, $df = 1$, $P < 0.001$; bird 157: $G = 3.68$, $df = 1$, $P = 0.06$, Fig. 4 Test 2). There were no significant differences in responses to different stimulus categories within ungrammatical set (all $G < \text{critical value } 5.99$, $df = 2$). Additional testing (Test 2+) indicated that birds had more difficulty to discriminate grammatical and ungrammatical stimuli when they were tested with the same non-adjacent dependencies but now occurring at arbitrary positions in the sequences. No significant differences were found in responding to

different stimuli sets in bird 150 ($G = 0.20$, $df = 1$, $P = 0.65$) and bird 157 ($G = 1.96$, $df = 1$, $P = 0.16$), but again bird 143 showed more responses to grammatical stimuli than to the ungrammatical stimuli ($G = 3.99$, $df = 1$, $P = 0.046$, appendix Fig. 1).

Phase 3: The performance during Training 3 indicated that all birds discriminated grammatical and ungrammatical stimuli right from the start and improved rapidly after this (Fig. 3 Training 3, appendix Table 3). When the intervening element type was changed from I_1 to I_2 (Test 3), two birds were still able to discriminate grammatical stimuli from ungrammatical ones (bird 150: $G = 91.94$, $df = 1$, $P < 0.001$; bird 143: $G = 109.82$, $df = 1$, $P < 0.001$; bird 157: $G = 2.34$, $df = 1$, $P = 0.13$, Fig. 4 Test 3). Within the ungrammatical set of Test 3, there were significant differences between responses to different categories of stimuli for two birds again (bird 150: $G = 9.33$, $df = 2$, $P = 0.01$; bird 143: $G = 11.97$, $df = 2$, $P < 0.01$; bird 157: $G = 1.02$, $df = 2$, $P = 0.60$). The STP G-test showed that bird 150 and 143, but not bird 157, had significantly more responses to Position-Correct stimuli (PC) than to the Co-Occurrence (CO) category while bird 143 also showed more responses to the category of Other stimuli. Although the responses to PC category were the highest for bird 150 and 143, they were still significantly lower than those to the grammatical set (bird 150: $G = 55.55$, $df = 1$, $P < 0.001$; bird 143: $G = 66.80$, $df = 1$, $P < 0.001$).

Phase 4: Birds discriminated grammatical and ungrammatical sets again better at the beginning of Training 4 compared to the performance they did at the beginning of Training 3 (Fig. 3, appendix Table 3). When switched to Test 4, in which the intervening items was changed from I_2 to I_3 , birds still responded significantly more to grammatical stimuli than to ungrammatical stimuli (all $G > \text{critical value } 3.84$, $df = 1$, all $P < 0.001$, Fig. 4 Test 4). Within the ungrammatical set of Test 4, there were also significant differences between responses to different categories of stimuli (bird 150: $G = 34.84$, $df = 2$, $P < 0.001$; bird 143: $G = 23.77$, $df = 2$, $P < 0.001$; bird 157: $G = 19.49$, $df = 2$, $P < 0.001$). Responses to PC were significantly higher than responses to the other categories for bird 150 and 143. However, the responses to PC were again lower than responses to the grammatical set (bird 150: $G = 67.28$, $df = 1$, $P < 0.001$; bird 143: $G = 107.29$, $df = 1$, $P < 0.001$). Responses to the three categories for bird 157 were significantly different from each other with most responses to Other stimuli. Overall, the birds seemed to pay more attention to the edge positions of the items since their responses to the PC category were highest in comparison with other categories within the ungrammatical set. However, this did not hinder their discrimination between grammatical and ungrammatical stimuli.

Phase 5: The learning curves of the grammatical and ungrammatical stimuli were distinct right from the beginning of Training 5 (Fig. 3 and appendix Table 3; as the birds reached training criterion or performed consistently from the beginning of Training 5 onwards, training did not continue till 4800 trials). Together with the other training and tests, the results of this experiment indicate the birds detected the non-adjacent dependencies even when both the distances between the paired items increased (from Training 2 to Training 3, number of intervening items increased from 1 to 2 and 3) and the intervening items changed (from the first

three training phases to Training 4 and 5, the intervening elements were changed from type 1 to type 2 and 3).

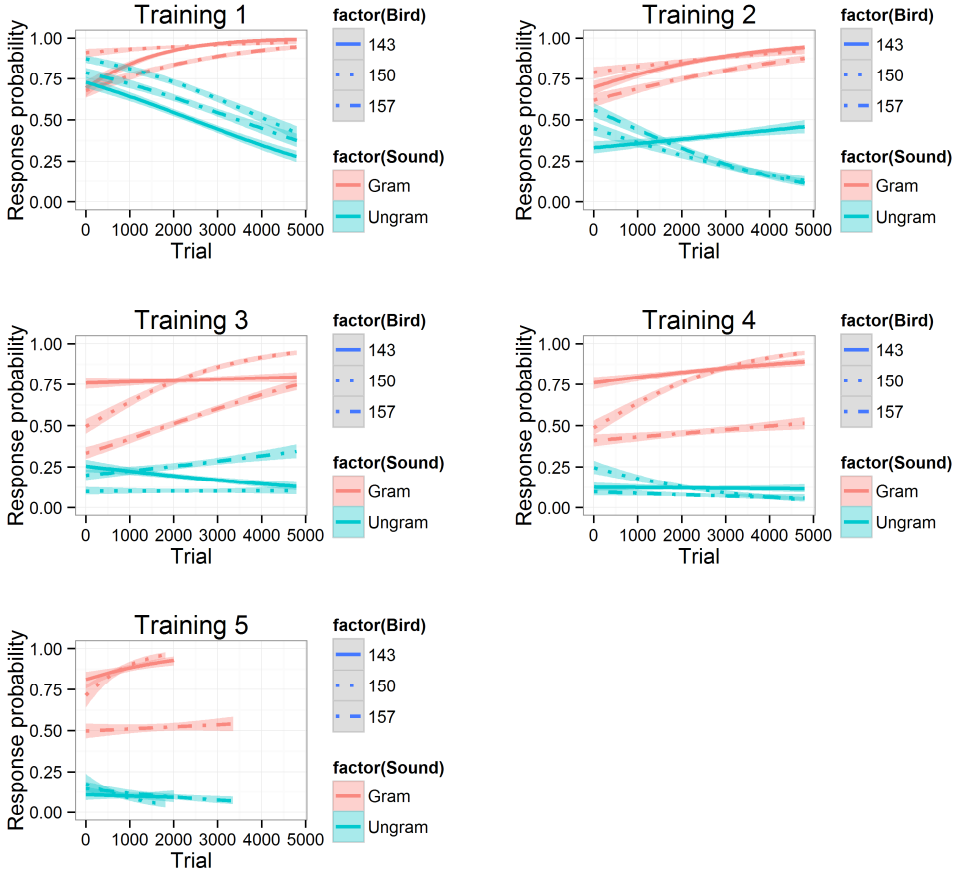


Fig. 3 Learning curves of the individual birds in Training 1-5, showing how the response to grammatical (indicated as Gram) and ungrammatical stimuli (indicated as Ungram) progressed over trials. Different line patterns and color indicate different individuals and sets of stimuli respectively. Broader color bands indicate the 95% confidence interval within one set.

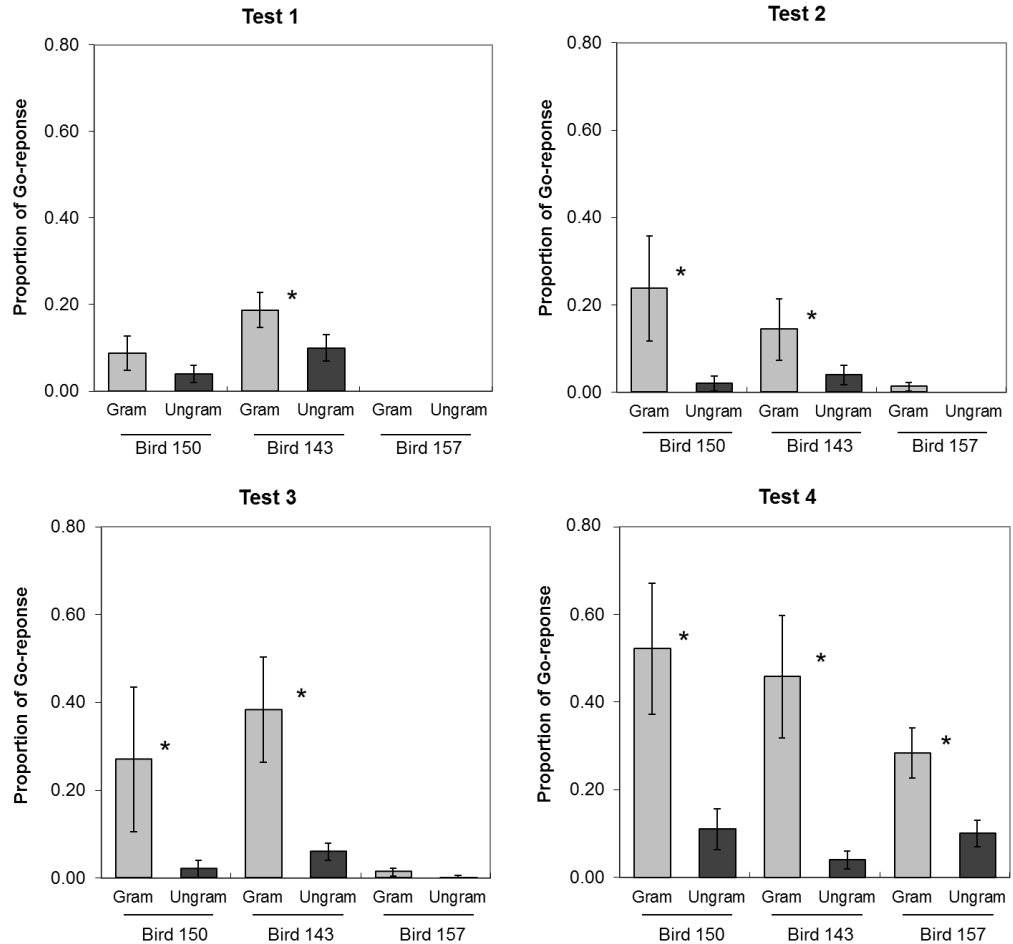


Fig. 4 Proportion of Go-responses of individual birds to grammatical (Gram) and ungrammatical stimuli (Ungram). A * indicates a significant difference between responses to grammatical and ungrammatical sets. In test 1, birds were exposed to test stimuli with one intervening item, such as $A I_1 B$. In test 2, birds were exposed to test stimuli with 2 to 3 intervening items, like $A I_1 I_1 B$ and $A I_1 I_1 I_1 B$. In test 3 and 4, test stimuli contained novel intervening items (I_2 and I_3 respectively).

Experiment 2. Non-adjacent dependencies at arbitrary positions

Phase 1: The 1st training was identical to that of experiment 1, with ‘starting small’ from adjacent dependencies. The one bird that participated in this experiment took a relatively long time to learn to discriminate between grammatical and ungrammatical stimuli (Fig. 5 Training1, appendix Table 4). But when tested with one intervening item inserted into the bigrams after the bird reaching training criterion, the bird responded significantly more to grammatical stimuli than to ungrammatical stimuli (Test 1: $G = 7.60$, $df = 1$, $P < 0.01$, Fig. 6).

Phase 2: Discrimination between grammatical and ungrammatical sets of stimuli was learned much faster during the 2nd training (Fig. 5 Training 2, appendix Table 4), although the dependencies were changed from adjacent to non-adjacent. In the following test (Test 2), non-adjacent dependencies between items were reallocated from items present at the edges to items

present at arbitrary positions in the sequences, the bird showed no significant difference in responding to the two sets of stimuli, although responses to ungrammatical stimuli were slightly higher than those to ungrammatical stimuli ($G = 0.72$, $df = 1$, $P = 0.40$, Fig. 6).

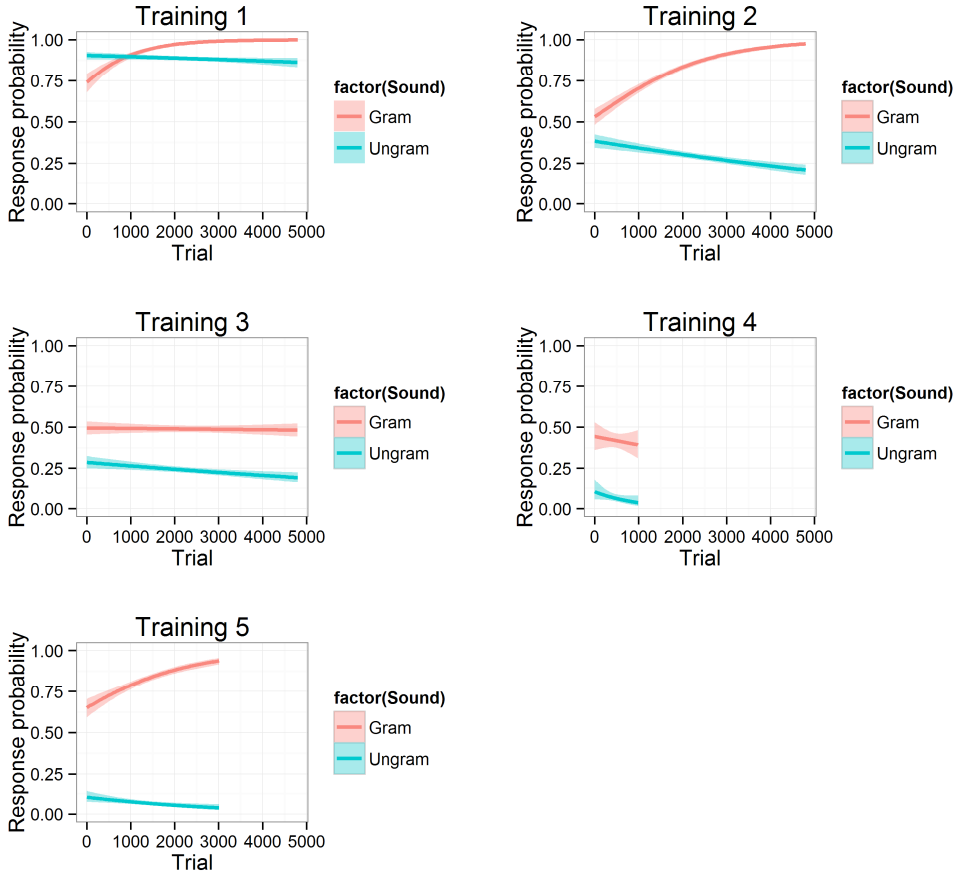


Fig. 5 Learning curves of bird 163 in Training 1-5, showing how response to grammatical (indicated as Gram) and ungrammatical stimuli (indicated as Ungram) progressed over trials. Different colors indicate different sets of stimuli. Broader color bands indicate the 95% confidence interval within one set.

Phase 3: Although the bird did not respond in a significantly different way to grammatical and ungrammatical sets during Test 2, it showed clear discrimination from the beginning of the 3rd training (Fig. 5 Training 3, appendix Table 4). In the 3rd test, some items from stimuli of the 3rd training were used but combined differently to generate novel test sequences. The bird discriminated the two stimulus sets perfectly ($G = 170.94$, $df = 1$, $P < 0.001$, Fig. 6). A more detailed test comparing responses to different categories of stimuli within ungrammatical set in Test 3 showed that there was no significant difference between responses to CO and Other stimuli ($G = 1.87$, $df = 2$, $P = 0.17$).

One more test (Test 4), with novel intervening items (I_1 replaced with I_2), was given before the bird was switched to the next training. In this test, the grammatical and ungrammatical stimuli were still discriminated significantly ($G = 60.40$, $df = 1$, $P < 0.001$, Fig. 6).

Phase 4: The stimuli in Training 4 (Fig. 5, appendix Table 4) were identical to stimuli of Test 4. Unfortunately, part of the learning data was lost due to a mechanical malfunction of the apparatus. Training 4 was stopped when we found the malfunction and bird was switched to the next phase soon afterwards. In the following test (Test 5), the intervening items of the stimuli were again replaced with another novel type (I_2 to I_3). The bird did not actively respond to these new test stimuli and showed no significant discrimination between the two sets ($G = 1.94$, $df = 1$, $P = 0.16$, Fig. 6)

Phase 5: However, in Training 5, using the same stimuli as in Test 5, the bird distinguished the grammatical and ungrammatical stimuli from the beginning of the training (Fig. 5, the bird reached training criterion at the beginning of Training 5, training did not continue till 4800 trials).

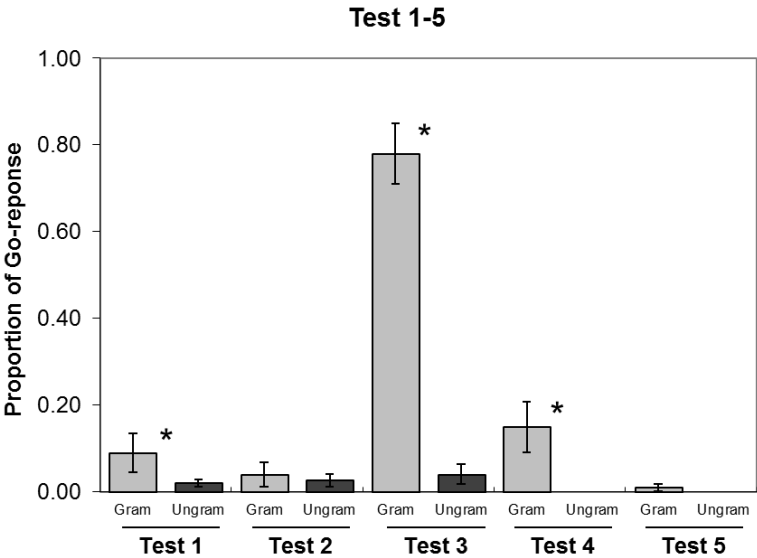


Fig. 6 Proportion of Go-responses of bird 163 to grammatical (Gram) and ungrammatical stimuli (Ungram). A * indicates a significant difference between responses to grammatical and ungrammatical sets. In test 1, the bird was exposed to test stimuli with one intervening item, such as $A I_1 B$. In test 2, the bird was exposed to 5-element strings with dependencies occurring at arbitrary positions of the strings, like $A I_1 B I_1 I_1$ and $I_1 A I_1 B I_1$. In test 3, the stimuli consisted of novel element sequences that never occurred in previous training. In test 4 and 5, the test stimuli contained novel intervening items (I_2 and I_3 respectively).

Discussion

Both experiments of the current study suggest that zebra finches are able to recognize the relation between two items, even when they are separated by intervening elements. One bird (143) already differentiated the grammatical and the ungrammatical pattern from each other after

the first training and after the first two training phases two birds were able to detect the non-adjacent dependencies over a distance of 2-3 intervening elements and they maintained this later on, when the type of intervening items between the dependent items was changed. It was not only the case for Experiment 1 in which the birds detected the non-adjacent dependencies when the dependent items occurred at the edges, but also for Experiment 2, where, after the first training phases, the bird detected non-adjacent dependencies at arbitrary positions in a sequence.

The training data show that birds transferred the knowledge about the connection between A and B, and C and D items of the stimuli from one phase to another. The speeding up of the learning process from phase 1 to 5 in both experiments is unlikely to be explained by familiarization with the task. If familiarization was the case, we would expect the learning curves to show a steep slope and the intercept (the starting points of both curves in the figures) of the grammatical and ungrammatical curves would not differ from each other. In the tests, the results seem most consistent with the birds having learned the non-adjacent dependencies at the edges. Nevertheless, the test results from Experiment 2 show some inconsistencies, in particular the non-significant differences between responses to grammatical and ungrammatical stimuli in tests 2 and 5. These were combined with a low response rate and suggest that the bird might have been reluctant to respond due to the changes of the sounds of the intervening I-elements. During the tests, birds were exposed to 20% test stimuli while they still received 80% reinforced training stimuli, which means that the bird may still get a sufficient number of rewards by being cautious and avoiding novel sounds. Although the bird did not show significant differences between the number of responses to grammatical and ungrammatical stimuli in every test, responses to grammatical stimuli were always higher than responses to ungrammatical sets.

The analysis of differences in responses within the ungrammatical set showed that birds responded more often to sequences with edge positions encoded correctly. This suggests that, similar to results from serial recall experiments (Chen, Swartz et al., 1997; Orlov, Yakovlev et al., 2000), the responses in Experiment 1 might be constrained by a positional memory: the birds linked items to their position in the sequence. Zebra finches also use positional information to detect differences between strings of song elements (Chen, van Rossum & ten Cate, in press). A higher response rate to an item in the correct position, but in the wrong combination can thus be considered a serial order intrusion error. The position based learning of items can also explain why Test 2+ of Experiment 1 suggested that detecting non-adjacent dependencies at novel, arbitrary, positions was more challenging than learning dependencies at the edges. It should, however, be noted that position based learning alone cannot explain the present data: a sequence starting with A can be either grammatical or ungrammatical, depending on the nature of the final element. Similarly a sequence ending with B is only correct when it has started with A. Hence the bird must make a connection between the nature of the first as well as that of the final element before it can correctly identify the sequence as grammatical. The observed discrimination between grammatical and ungrammatical stimuli can also not be explained by certain adjacent transitions or element co-occurrence. Zebra finches can readily detect such adjacent dependencies, as shown in several studies (van Heijningen, Chen et al., 2013; van Heijningen, de Visser et al., 2009; Chen & ten Cate, submitted). However, in the current experiment the relevant A, B, C and D elements were always separated by I-elements and hence

a combination like AI or IB does not provide a cue about whether it is grammatical or ungrammatical. Also, the number of transitions between different types of elements present in a sequence, which might have been a cue used in the experiments of Ravignani, Sonnweber et al. (2013) does not differ between grammatical and ungrammatical strings in the present experiment.

So, we conclude that our findings show that the ability to detect the relation between two items, even when they are non-adjacent in a sequence, also exists in a songbird species. Zebra finches were not only able to learn non-adjacent dependencies at an arbitrary distance but also can learn non-adjacent dependencies at arbitrary positions of a sequence, which so far has not been shown in any other non-human animal. Our results, in combination with those obtained earlier for tamarins (Newport, Hauser et al., 2004) and squirrel monkeys (Ravignani, Sonnweber et al., 2013) indicate that non-human animals may detect at least some types of non-adjacent dependencies in acoustic stimuli, partly bridging the gap between the seemingly quite different computational abilities of humans and other animals. The data for tamarins suggest a stimulus-dependent difference between tamarins and humans in whether or not they can detect the pattern, and hence that the nature of the cognitive mechanism involved may differ between different species. Nevertheless these studies as well as the current one indicate that the ability to detect relations between non-adjacent items in vocal sequences is not linked to having language or to whether or not such dependencies are present in the natural vocalizations of animals, as there is no evidence of the presence of such patterns in monkey calls or birdsongs. It suggests that the ability to detect and use such relations in human language may have originated from co-opting a cognitive ability from outside the domain of language.

References

- Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences*, 15(3), 113-121. doi: 10.1016/j.tics.2011.01.002
- Bolhuis, J. J., & Everaert, M. (2013). *Birdsong, speech, and language: exploring the evolution of mind and brain*: MIT press.
- Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, 11(11), 747-759. doi: 10.1038/Nm2931
- Chen, S., Swartz, K. B., & Terrace, H. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science*, 8(2), 80-86.
- Cochran, B. P., McDonald, J. L., & Parault, S. J. (1999). Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language*, 41(1), 30-58.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. doi: 10.1146/annurev.neuro.22.1.567
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1), 71-99.
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal Cognition*, 13(3), 483-495. doi: 10.1007/s10071-009-0299-8
- Endress, A. D., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, 13(8), 348-353. doi: dx.doi.org/10.1016/j.tics.2009.05.005
- Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics*, 25(4), 166-177.
- Gabriel, K. R. (1966). Simultaneous Test Procedures for Multiple Comparisons on Categorical Data. *Journal of the American Statistical Association*, 61(316), 1081-&. doi: 10.2307/2283202
- Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, 13(5), 431-436.
- Gómez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7(2), 183-206.
- 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(5598), 1569-1579.
- 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(3), B53-64.
- Henson, R. N. (1998). Short-term memory for serial order: The start-end model. *Cogn Psychol*, 36(2), 73-137.
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., . . . Shimizu, T. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6(2), 151-159.
- Lachlan, R., Verhagen, L., Peters, S., & Cate, C. t. (2010). Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taenopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *Journal of Comparative Psychology*, 124(1), 92.
- Lai, J. (2013). *The learnability of center-embedded recursion: experimental studies with artificial and natural language*: Department of Cognitive Psychology, Faculty of Social and Behavioural Sciences, Leiden University.
- Lany, J., & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science*, 19(12), 1247-1252.
- Lany, J., Gómez, R. L., & Gerken, L. A. (2007). The role of prior experience in language acquisition. *Cognitive Science*, 31(3), 481-507.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cogn Psychol*, 48(2), 127-162.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cogn Psychol*, 49(2), 85-117.
- Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, 404(6773), 77-80.
- Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: no need for algebraic-like computations. *Journal of Experimental Psychology: General*, 133(4), 573.
- Ravignani, A., Sonnweber, R.-S., Stobbe, N., & Fitch, W. T. (2013). Action at a distance: dependency sensitivity in a New World primate. *Biology letters*, 9(6), 20130852.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, 69(2), 105-134.

- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: the principles and practice of statistics in biological research*. New York: Freeman and Co.
- 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(4), 706.
- Toro, J. M., & Trobalon, J. B. (2005). Statistical computations over a speech stream in a rodent. *Percept Psychophys*, 67(5), 867-875.
- 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(2), 165-175. doi: 10.1007/s10071-012-0559-x
- 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. *Proceeding of the National Academy of Sciences of the United States of America*, 106(48), 20538-20543. doi: 10.1073/pnas.0908113106
- Zann, R. A., & Bamford, M. (1996). *The zebra finch: a synthesis of field and laboratory studies*.: Oxford University Press Oxford.

Appendix

Table 1 Training and Test stimuli of Experiment 1

Training			Test		Additional Test		
	Go	Nogo					
1	AB	AD	1	AI ₁ B	AI ₁ D		
	CD	AC BD CB CA DB		CI ₁ D	CI ₁ B AI ₁ C BI ₁ D CI ₁ A DI ₁ B		
2	AI ₁ B	AI ₁ D	2	AI ₁ I ₁ B	AI ₁ I ₁ D	2+	AI ₁ BI ₁ I ₁ I ₁ AI ₁ BI ₁ CI ₁ DI ₁ I ₁ I ₁ CI ₁ I ₁ D
	CI ₁ D	CI ₁ B AI ₁ C BI ₁ D CI ₁ A DI ₁ B		AI ₁ I ₁ I ₁ B CI ₁ I ₁ D CI ₁ I ₁ I ₁ D	AI ₁ I ₁ I ₁ D CI ₁ I ₁ B CI ₁ I ₁ I ₁ B BI ₁ I ₁ D DI ₁ I ₁ C		
3	AI ₁ I ₁ B	AI ₁ I ₁ D	3	AI ₂ B	AI ₂ D		
	AI ₁ I ₁ I ₁ B CI ₁ I ₁ D CI ₁ I ₁ I ₁ D	AI ₁ I ₁ I ₁ D CI ₁ I ₁ B CI ₁ I ₁ I ₁ B BI ₁ I ₁ D DI ₁ I ₁ C		AI ₂ I ₂ B AI ₂ I ₂ I ₂ B CI ₂ D CI ₂ I ₂ D CI ₂ I ₂ I ₂ D	AI ₂ I ₂ D AI ₂ I ₂ I ₂ D AI ₂ I ₂ I ₂ B CI ₂ B CI ₂ I ₂ B CI ₂ I ₂ I ₂ B BI ₂ I ₂ A DI ₂ I ₂ C BI ₂ I ₂ D CI ₂ I ₂ A		
4	AI ₂ B	AI ₂ D	4	AI ₃ B	AI ₃ D		
	AI ₂ I ₂ B AI ₂ I ₂ I ₂ B CI ₂ D CI ₂ I ₂ D CI ₂ I ₂ I ₂ D	AI ₂ I ₂ D AI ₂ I ₂ I ₂ D CI ₂ B CI ₂ I ₂ B CI ₂ I ₂ I ₂ B DI ₂ I ₂ C BI ₂ I ₂ D		AI ₃ I ₃ B AI ₃ I ₃ I ₃ B CI ₃ D CI ₃ I ₃ D CI ₃ I ₃ I ₃ D	AI ₃ I ₃ D AI ₃ I ₃ I ₃ D AI ₃ I ₃ I ₃ B CI ₃ B CI ₃ I ₃ B CI ₃ I ₃ I ₃ B BI ₃ I ₃ A DI ₃ I ₃ C BI ₃ I ₃ D CI ₃ I ₃ A		
5	AI ₃ B	AI ₃ D					
	AI ₃ I ₃ B AI ₃ I ₃ I ₃ B CI ₃ D CI ₃ I ₃ D CI ₃ I ₃ I ₃ D	AI ₃ I ₃ D AI ₃ I ₃ I ₃ D CI ₃ B CI ₃ I ₃ B CI ₃ I ₃ I ₃ B DI ₃ I ₃ C BI ₃ I ₃ D					

The letters A, B, C, D and I indicate 5 different elements. Elements A, B, C and D were used to generate non-adjacent dependencies while element I was used as the intervening item. The numbers 1 to 3 indicate different types of intervening items.

Table 2. Training and Test stimuli of Experiment 2

Training			Test		
	Go	Nogo			
1	AB	AD	1	AI ₁ B	AI ₁ C
	CD	AC BD CB CA DB		CI ₁ D	AI ₁ D BI ₁ D CI ₁ A CI ₁ B DI ₁ B
2	AI ₁ B	AI ₁ D	2	AI ₁ I ₁ I ₁ B	AI ₁ I ₁ I ₁ D
	CI ₁ D	AI ₁ C BI ₁ D CI ₁ B CI ₁ A DI ₁ B		AI ₁ BI ₁ I ₁ I ₁ AI ₁ BI ₁ CI ₁ DI ₁ I ₁ I ₁ CI ₁ I ₁ D	I ₁ AI ₁ DI ₁ I ₁ I ₁ DI ₁ B CI ₁ BI ₁ I ₁ I ₁ CI ₁ I ₁ B
3	AI ₁ I ₁ I ₁ B AI ₁ BI ₁ I ₁ I ₁ AI ₁ BI ₁ I ₁ I ₁ AI ₁ B CI ₁ I ₁ DI ₁ CI ₁ DI ₁ I ₁ I ₁ CI ₁ I ₁ D I ₁ I ₁ CI ₁ D	AI ₁ I ₁ I ₁ D I ₁ AI ₁ DI ₁ CI ₁ BI ₁ I ₁ I ₁ CI ₁ I ₁ B AI ₁ CI ₁ I ₁ I ₁ I ₁ DI ₁ B CI ₁ I ₁ AI ₁ I ₁ I ₁ BI ₁ D	3	AI ₁ I ₁ BI ₁ I ₁ AI ₁ I ₁ B CI ₁ I ₁ I ₁ D I ₁ CI ₁ DI ₁	AI ₁ I ₁ DI ₁ I ₁ CI ₁ BI ₁ I ₁ BI ₁ AI ₁ I ₁ DI ₁ CI ₁ I ₁ BI ₁ I ₁ D I ₁ I ₁ AI ₁ C I ₁ BI ₁ CI ₁ I ₁ CI ₁ AI ₁ DI ₁ I ₁ BI ₁ I ₁ DI ₁ AI ₁
				4	AI ₂ I ₂ I ₂ B I ₂ AI ₂ BI ₂ CI ₂ I ₂ DI ₂ I ₂ CI ₂ I ₂ D AI ₂ I ₂ BI ₂ I ₂ AI ₂ I ₂ B CI ₂ I ₂ I ₂ D I ₂ CI ₂ DI ₂
4	AI ₂ I ₂ I ₂ B I ₂ AI ₂ BI ₂ CI ₂ I ₂ DI ₂ I ₂ CI ₂ I ₂ D AI ₂ I ₂ BI ₂ I ₂ AI ₂ I ₂ B CI ₂ I ₂ I ₂ D I ₂ CI ₂ DI ₂	AI ₂ I ₂ I ₂ D I ₂ AI ₂ DI ₂ CI ₂ BI ₂ I ₂ I ₂ CI ₂ I ₂ B AI ₂ I ₂ DI ₂ I ₂ CI ₂ BI ₂ CI ₂ I ₂ I ₂ B I ₂ BI ₂ I ₂ D	5	AI ₃ I ₃ I ₃ B I ₃ AI ₃ BI ₃ CI ₃ I ₃ DI ₃ I ₃ CI ₃ I ₃ D AI ₃ I ₃ BI ₃ I ₃ AI ₃ I ₃ B CI ₃ I ₃ I ₃ D I ₃ CI ₃ DI ₃	AI ₃ I ₃ I ₃ D I ₃ AI ₃ DI ₃ CI ₃ BI ₃ I ₃ I ₃ CI ₃ I ₃ B AI ₃ I ₃ DI ₃ I ₃ CI ₃ BI ₃ CI ₃ I ₃ I ₃ B I ₃ BI ₃ I ₃ D
5	AI ₃ I ₃ I ₃ B I ₃ AI ₃ BI ₃ CI ₃ I ₃ DI ₃ I ₃ CI ₃ I ₃ D AI ₃ I ₃ BI ₃ I ₃ AI ₃ I ₃ B CI ₃ I ₃ I ₃ D I ₃ CI ₃ DI ₃	AI ₃ I ₃ I ₃ D I ₃ AI ₃ DI ₃ CI ₃ BI ₃ I ₃ I ₃ CI ₃ I ₃ B AI ₃ I ₃ DI ₃ I ₃ CI ₃ BI ₃ CI ₃ I ₃ I ₃ B I ₃ BI ₃ I ₃ D			

The letters A, B, C, D and I indicate 5 different elements. Elements A, B, C and D were used to generate non-adjacent dependencies while element I was used as the intervening item. The numbers 1 to 3 indicate different types of intervening items.

Table 3 Statistical models and parameter estimates describing the relationship between response probability and training trials for different stimulus types.

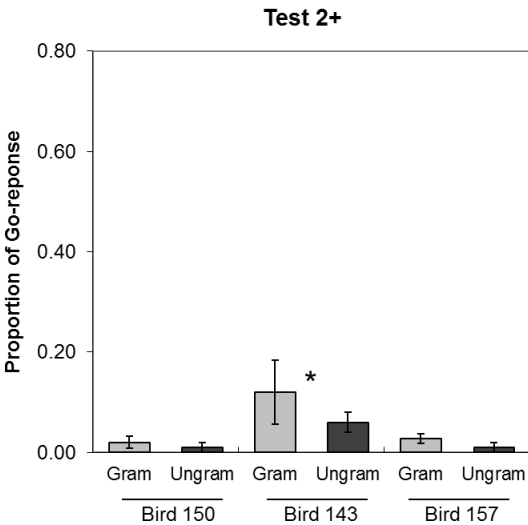
Training	Stimulus type & bird	Fitted Model
1	Ungram 143	Logit[Response]= 0.98518** - 0.000405** Trial
	Ungram 150	Logit[Response]= 1.927247** - 0.000468** Trial
	Ungram 157	Logit[Response]= 1.312305** - 0.00038** Trial
	Gram 143	Logit[Response]= 0.805128** + 0.000853** Trial
	Gram 150	Logit[Response]= 2.324421** + 0.000283** Trial
	Gram 157	Logit[Response]= 0.73903** + 0.000445** Trial
2	Ungram 143	Logit[Response]= -0.701174** + 0.000111** Trial
	Ungram 150	Logit[Response]= -0.21224* - 0.000351** Trial
	Ungram 157	Logit[Response]= 0.239915** - 0.000478** Trial
	Gram 143	Logit[Response]= 0.832201** + 0.000416** Trial
	Gram 150	Logit[Response]= 1.313264** + 0.000242** Trial
	Gram 157	Logit[Response]= 0.482761** + 0.000305** Trial
3	Ungram 143	Logit[Response]= -1.070476** - 0.000174** Trial
	Ungram 150	Logit[Response]= -2.202433** + 8e-06 Trial
	Ungram 157	Logit[Response]= -1.388957** + 0.000155** Trial
	Gram 143	Logit[Response]= 1.149216** + 4.4e-05 Trial
	Gram 150	Logit[Response]= -0.021112 + 0.000606** Trial
	Gram 157	Logit[Response]= -0.694645** + 0.000371** Trial
4	Ungram 143	Logit[Response]= -1.942141** - 2.1e-05 Trial
	Ungram 150	Logit[Response]= -1.119934** - 0.000395** Trial
	Ungram 157	Logit[Response]= -2.23171** - 0.000116 Trial
	Gram 143	Logit[Response]= 1.156997** + 0.000196** Trial
	Gram 150	Logit[Response]= -0.052115 + 0.000609** Trial
	Gram 157	Logit[Response]= -0.36397** + 8.9e-05** Trial
5	Ungram 143	Logit[Response]= -2.10415** - 9.4e-05 Trial
	Ungram 150	Logit[Response]= -1.54446** - 8e-04** Trial
	Ungram 157	Logit[Response]= -1.764913** - 0.000251** Trial
	Gram 143	Logit[Response]= 1.455258** + 0.000557** Trial
	Gram 150	Logit[Response]= 0.908288** + 0.001287** Trial
	Gram 157	Logit[Response]= -0.011014** + 5.1e-05 Trial

Logit[] indicates binomial regression using the logit link function, i.e., $\text{Logit}(Y) = \ln \frac{y}{1-y}$. *, p<0.05, **, p<0.01.

Table 4 Statistical models and parameter estimates describing the relationship between response probability and training trials for different stimulus types to bird 163.

Training	Stimulus type	Fitted Model
1	Ungram	Logit[Response]= 2.252743** - 8.6e-05 Trial
	Gram	Logit[Response]= 1.043491** + 0.001238** Trial
2	Ungram	Logit[Response]= -0.476072** - 0.000178** Trial
	Gram	Logit[Response]= 0.120042** + 0.000745** Trial
3	Ungram	Logit[Response]= -0.915653** - 0.000107** Trial
	Gram	Logit[Response]= -0.018636 - 1.1e-05 Trial
4	Ungram	Logit[Response]= -2.176687** - 0.001121** Trial
	Gram	Logit[Response]= -0.22574 - 0.000214 Trial
5	Ungram	Logit[Response]= -2.154493** - 0.000352** Trial
	Gram	Logit[Response]= 0.620598** + 0.000692** Trial

Logit[] indicates binomial regression using the logit link function, i.e., $\text{Logit}(Y) = \ln \frac{y}{1-y}$. *, p<0.05, **, p<0.01.



Appendix Fig. 1 Proportion of Go-responses of individual birds to grammatical (Gram) and ungrammatical stimuli (Ungram) for test 2+. In this test, the birds were exposed to 5-element strings with dependencies occurring at arbitrary positions of the strings, such as AI₁BI₁I₁A. * indicates a significant difference between responses to grammatical and ungrammatical sets.

Chapter 6

Summary and conclusion

Language is considered a key component that distinguishes humans from other animals. However, since the 19th century, researchers started to realize that some features related to language and language learning were shared with other animals. The debates on to what extent the language faculty is shared with nonhuman animals has boosted an increasing number of cognitive studies with animals. This thesis contributes to such cognitive research by examining whether learning abilities that are considered important in human language learning are also present in zebra finches. This can indicate that these processes may have a domain general origin and, during evolution, may have been co-opted for language.

Chapter 2 addressed the question: what mechanisms are involved in learning a sequence of vocal items in zebra finches. In this study, the birds were trained with sequences differing in transitional and positional information and subsequently tested with novel strings containing positional and transitional similarities with the training strings. When tested with two strings that shared more transitional than positional information with the training strings, birds were more likely to use transitional than positional information to recognize the test stimuli. When the test stimuli contained more positional similarity with one training stimulus and more transitional similarity with the other training stimulus, preferences for using positional or transitional information were both found in different birds, which suggested that both encoding mechanisms were involved in the discrimination. Unlike a previous study on sequential learning in a songbird, the starling (Comins & Gentner, 2010), which suggested that songbirds had a learning bias for positional information in the encoding of sequences, the results of this chapter suggested that the zebra finches are able to bias their sequential coding strategy depending on what might be most profitable given the context.

This study also showed that zebra finches were able to discriminate segments of the training strings according to their chaining regularities, even though the length of segments was much shorter than that of the training strings. A fundamental process during language acquisition is to segment words from fluent speech. Humans can accomplish this task by paying attention to transitional cues in continuous speech. The ability of zebra finch to employ transitional cues indicates that some similarities might be shared between the cognitive mechanisms of this bird species and humans.

Chapter 3 addressed whether zebra finches are able to discriminate between, and generalize, affixation patterns. Zebra finches were trained to discriminate artificial sequences resembling prefixed and suffixed ‘words’. The ‘stems’ of the ‘words’, consisted of different combinations of a triplet of song elements, to which a fourth element was added that acted as either a prefix or a suffix. Six zebra finches received the prefixed words as Go stimuli and the suffixed ones as Nogo stimuli, for another six birds the conditions were reversed. The birds were first trained with a set of affixed stems and subsequently tested with novel stems, which were composed either by rearranging the training stems (familiar element types) or by using novel element types.

The results show that the birds were able to generalize the affixation patterns to novel stems, with both familiar element types and novel element types. It indicates that the discrimination resulting from the training was not based on memorization of individual stimuli,

but on a shared property among Go or Nogo stimuli, i.e. their affixation patterns. We found that the birds were able to learn both regularities. Remarkably, it seemed that using the suffix as a cue was much weakened when birds trained with prefixation as the Go pattern. On the other hand, birds were trained with suffixation as the Go pattern showed clear evidence of using both prefix and suffix. This may imply an interesting parallel to the asymmetry in the type of affixation preferred in human languages.

Chapter 4 deals with a controversial topic that is recently getting a lot of attention. Discovering structured rules is a hallmark of human linguistic abilities. To demonstrate the acquisition of abstract rules, successful categorization should not depend on the physical similarity of strings of items that are all structured according to the same algorithm, but should also be possible when the items of the individual strings bear no acoustic, but only structural similarity to each other. Currently, few studies have addressed the question whether the animals are able to generalize structures without relying on the physical similarities between stimuli as humans do.

In a series of discrimination experiments, we presented zebra finches and human adults with comparable training and tests with the same artificial stimuli consisting of *XYX* and *XXY* structures, in which *X* and *Y* were zebra finches song elements. Zebra finches readily discriminated the training stimuli. They showed evidence of simple rule abstraction related to positional learning. Some birds also discriminated novel stimuli when these were composed of familiar element types, but none of the birds generalized the patterns to novel element types. These findings suggest that zebra finches may be able to a stimulus bound generalization, but we found no evidence for a more abstract rule generalization. This differed from the human adults, who readily categorized test stimuli consisting of novel element types into different groups according to their structure. The limited abilities for rule abstraction in zebra finches may indicate that the precursor of more complex abstractions is a stimulus bound rule learning mechanism (that might have become more abstract under pressure of the growing lexicon in human language).

Chapter 5 examined the ability of zebra finches to learn nonadjacent dependencies between items in a string of vocal elements. The birds were trained and tested with an artificial grammar consisting of different nonadjacent dependencies. Zebra finch song elements were arranged in strings according to a particular grammar involving a non-adjacent dependency between two items. In experiment 1 of this chapter, the two elements involved in the dependency always occurred at the edges of sequences but the number and nature of intervening elements varied, creating strings with varied length. One step further, experiment 2 tested the ability of this songbird species to detect dependencies between elements that occurred at arbitrary positions in the sequences.

Results from Experiment 1 showed that birds generalized nonadjacent dependencies over varied distances. They were able to discriminate novel stimuli according to the relations between non-adjacent items even when the intervening items were never heard by the subjects before. Nonadjacent dependencies in human language do not always occur at the edges, they can occur at arbitrary positions in speech sequences. In the second experiment, one zebra finch

learned to detect non-adjacent dependencies in different positions in the stimulus sequences. The results of this chapter suggest that the quite complex cognitive capacity to detect relationships between elements that are non-adjacent in a vocal sequence is present in zebra finches.

Conclusion

To conclude, this thesis showed that songbirds are not only able to encode sequences by using positional cues, but also to use transitional information between items for sequential encoding. The clear presence of using transitional information in the zebra finches is of particular interest from a comparative perspective. Transitional information has been shown in an experimental setting to be exploited by human infants in a fundamental process of language acquisition, in which words were segmented from continuous speech (Saffran, Aslin et al., 1996). In addition to previous findings from primates and rats (e.g. Hauser, Newport et al., 2001; Toro & Trobalon, 2005), birds now are also shown to be sensitive to transitional information, suggesting that such a cognitive skill is broadly present and may have been recruited for language learning from a more general domain. Our findings are by no means suggesting that positional cues are less important. On the contrary, we suggest that learning of the structure of a sequence might well be constrained by perceptual and memory primitives, such as seen in the learning of the edge-based positional regularities. Zebra finches in this thesis were shown to generalize edge-based positional patterns resembling the surface transformations of human prefixation and suffixation. This suggests that affixation patterns in human languages might have a prelinguistic origin, perhaps related to serial position learning and co-opted for use in language.

Of course, there are more sophisticated regularities that do not simply rely on the positions of items or linear transitions between items. To generalize a more abstract rule, one needs to make categorizations without basing this on the physical similarity between items. In this thesis, zebra finches were found to discriminate an *XYX* from an *XXY* structure with familiar stimuli only. Together with previous studies, which provided no strong evidence for such rule generalization (Hauser & Glynn, 2009; Murphy, Mondragon et al., 2008; Seki, Suzuki et al., 2013; Toro & Trobalon, 2005; van Heijningen, Chen et al., 2013, discussed in chapter 4), this divergence from what has been observed in humans suggests that the ability to generalize ‘algebraic’ rules might be unique to humans. However, more studies are still needed to make sure that the failure of finding this ability is not due to methodological factors. Finally, in this thesis, another more sophisticated regularity, that of detecting nonadjacent dependency relations between items, was shown to be present in the zebra finches. The birds were able to learn nonadjacent dependencies at varied distance and at arbitrary positions of sequences. This latter ability has not been demonstrated in any non-human animal study before and thus provides a new perspective on what animals can do and what is shared with humans.

Altogether, this thesis provides positive evidence for similarities between humans and songbirds in using transitional information, generalizing surface transformations of human affixation patterns and detecting nonadjacent dependencies. The presence of such similarities between species that shared a common ancestor some 300 million years ago may indicate that either such abilities are rooted in a deep past, or, more likely, that they evolved independently

and, in the case of zebra finches, unrelated to their use in natural vocal communication. Demonstrating these similarities in the highly constrained and simplified conditions in our experiments cannot result in the conclusion that ‘birds can do the same thing as humans’. However, they do demonstrate cognitive abilities that may also have been present in prelingual humans and that may have been evolutionary precursors co-opted for constructing the complex syntax processing machinery present in modern humans.

Reference

- Comins, J. A., & Gentner, T. Q. (2010). Working memory for patterned sequences of auditory objects in a songbird. *Cognition*, 117(1), 38-53. doi: 10.1016/j.cognition.2010.06.009
- Hauser, M. D., & Glynn, D. (2009). Can free-ranging rhesus monkeys (*Macaca mulatta*) extract artificially created rules comprised of natural vocalizations? *Journal of Comparative Psychology*, 123(2), 161-167. doi: 10.1037/a0015584
- 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(3), B53-64.
- Murphy, R. A., Mondragon, E., & Murphy, V. A. (2008). Rule learning by rats. *Science*, 319(5871), 1849-1851. doi: 10.1126/science.1151564
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.
- Seki, Y., Suzuki, K., Osawa, A. M., & Okanoya, K. (2013). Songbirds and humans apply different strategies in a sound sequence discrimination task. *Frontiers in psychology*, 4, 447. doi: 10.3389/fpsyg.2013.00447
- Toro, J. M., & Trobalon, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & psychophysics*, 67(5), 867-875.
- 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(2), 165-175. doi: 10.1007/s10071-012-0559-x

Nederlandse samenvatting

(Dutch summary)

Een van de belangrijke verschillen tussen mensen en andere dieren is de complexiteit van de menselijke taal. Hoewel deze uniek is, is al sinds de 19e eeuw bekend dat aspecten van taal, en het leren ervan, ook aanwezig zijn in andere diersoorten. Het debat over de mate waarin “the language faculty” aanwezig is bij verschillende diersoorten, heeft een stimulans gegeven voor meer onderzoek naar de cognitie van dieren. Het werk in dit proefschrift draagt hieraan bij door te onderzoeken of leer vermogens die bij mensen belangrijk zijn voor taalontwikkeling, ook aanwezig zijn bij zebrevinken. Gedeelde vermogens tussen mensen en andere dieren kunnen een indicatie zijn dat deze processen een meer algemene oorsprong hebben en, door evolutie, verder ontwikkeld zijn voor menselijke taal.

Hoofdstuk 2 behandelt de vraag: welke mechanismen zijn betrokken bij het leren van een reeks zebrevink zangelementen? Tijdens dit experiment werden de vogels getraind met reeksen die onderscheiden konden worden door de positie van de elementen en in de mate waarin de elementen na elkaar voorkwamen. Hierna werden de zebrevinken getest met nieuwe reeksen die overeenkomsten hadden met de trainingsreeksen wat betreft de posities en het samen voorkomen van elementen. Afhankelijk van de testreeksen, gebruikten de zebrevinken de positie of het samen voorkomen van de elementen. Dit suggereert dat beide systemen betrokken waren bij het leren van reeksen geluiden. Dit resultaat is in tegenstelling tot een eerdere studie naar het leren van reeksen door spreeuwen, die suggereerde dat zangvogels een voorkeur hadden voor het gebruik van positionele informatie bij het leren van reeksen. Onze resultaten tonen aan dat de zebrevinken hun strategie kunnen aanpassen aan de beste optie in een bepaalde context. Daarnaast tonen ze ook aan dat zebrevinken, net als mensen, de overgangen tussen elementen kunnen gebruiken. Dit is een indicatie dat er mogelijk overeenkomsten zijn tussen de cognitieve mechanismen van zebrevinken en mensen.

Hoofdstuk 3 beschrijft onderzoek naar het maken van onderscheid tussen vervoegingspatronen, en het generaliseren van deze patronen, door zebrevinken. Zebrevinken werden getraind om onderscheid te maken tussen artificiële geluidsreeksen die in structuur vergelijkbaar waren met woorden met voor- en achtervoegsels. De zebrevinken werden getraind met ‘stam’, bestaande uit een variabel samengestelde set van 3 zangelementen, voorzien van een vaste vervoeging en werden daarna getest met vervoegingen met een onbekende stam. Deze stam kon bestaan uit een nieuwe combinatie van de stam elementen uit de training of uit nieuwe elementen. De resultaten laten zien dat de vogels in staat waren om zowel de voor- als achtervoegsels te generaliseren naar reeksen met een nieuwe stam, ongeacht of deze bestond uit bekende of nieuwe elementen. Dit geeft aan dat het onderscheid dat de vogels maakten niet gebaseerd was op herinnering van de afzonderlijke stimuli, maar op een gedeelde eigenschap van de stimuli, zoals het patroon van vervoegingen.

Hoofdstuk 4 onderzoekt het leren en abstraheren van structuurregels in geluidsreeksen; een controversieel onderwerp dat momenteel veel aandacht verkrijgt. Om aan te tonen dat een

abstracte representatie van de regels is gevormd, zou categorisering van nieuwe items niet afhankelijk moeten zijn van klankovereenkomsten tussen de items, maar van de structurele overeenkomsten. De bestaande studies geven geen duidelijk antwoord op de vraag of dieren in staat zijn om structurele regels te leren op een manier vergelijkbaar met hoe mensen dit doen. In een serie experimenten kregen zowel zebrovinken als volwassen mensen vergelijkbare trainingen en tests met artificiële stimuli die een *XYX* of een *XXY* structuur hadden. Zebrovinken maakten duidelijk onderscheid tussen de twee structuren in de training stimuli en konden de positie van de elementen generaliseren naar nieuwe combinaties. Sommige vogels konden ook correct onderscheid maken tussen nieuwe stimuli die uit bekende element types bestonden, maar geen enkele vogel kon dit met stimuli van onbekende elementen. Dit suggereert dat zebrovinken een generalisatie kunnen maken met bekende elementen, maar er was geen bewijs van generalisatie van de abstracte *XYX* en *XXY* regels. Volwassen mensen konden wel stimuli met nieuwe geluiden onderscheiden op basis van de structuur en generaliseerden dus de abstracte regels. De beperkte capaciteiten van zebrovinken voor het abstraheren van regels kan aangeven dat de voorganger van complexe abstracties een stimulusspecifieke generalisatie was.

Hoofdstuk 5 beschrijft het onderzoek naar het vermogen van zebrovinken om de afhankelijkheid te leren van twee elementen die niet direct na elkaar in een reeks voorkomen. De vogels werden getraind en getest met reeksen zebrovink zangelementen waarbij twee elementen altijd samen voorkwamen, maar nooit direct na elkaar. In experiment 1 stonden de elementen die bij elkaar hoorden altijd aan het begin en het einde van een reeks. Het aantal elementen dat daartussen voorkwam varieerde, waardoor er langere en kortere reeksen ontstonden. In experiment 2 komen de twee elementen op verschillende plekken in de reeks voor. Dit laatste experiment is meer vergelijkbaar met de situatie in menselijke spraak. Zebrovinken waren in beide experimenten in staat om de afhankelijkheid van de twee elementen te herkennen en te generaliseren naar nieuwe reeksen. Dit toont aan dat niet alleen mensen in staat zijn tot het volbrengen van een dergelijke complexe cognitieve taak.

Samengevat, laat dit proefschrift zien dat sommige capaciteiten die specifiek voor taal leken te zijn, ook, zij het in een simpele vorm, voorkomen bij andere diersoorten. Dit geeft inzicht in de cognitieve capaciteiten van waaruit menselijke taal geëvolueerd kan zijn.

Acknowledgements

First of all, my sincere gratitude goes to my supervisor Carel ten Cate for all the support and insightful scientific guidance through this research. Thank you for your encouragement and patience with me through this four-year doctoral study. I am grateful to our group members Hans, Katharina, Peter, Caroline, Sita, Wouter, Errol, Michelle, Sabine, Saeed, Pralle, Fleur, Ozkan and Sissy for providing such an inspiring and joyful environment for doing research, especially to Michelle, Pralle and Errol for many scientific and non-scientific interactions. Thank you all for your help during my study, your critical and helpful feedback and discussion of my research findings. Special thanks to Michelle for your help with the Dutch summary and to Pralle for the comments on the introduction of this thesis, also thank you for being my paranymphs. I thank the students Tessa, Naomi, Flora, Danielle, Sinan, Iris and Daphne, who helped me with my experiments. I also appreciated the help of the volunteers for participating in the human experiment in this thesis. I want to thank all of my friends in Leiden, though it is difficult to acknowledge everyone here: thank you for helping me and giving me mental support along this journey. I want to thank Chuan for being there for me in many ways and also for your help with the statistics in this thesis. Last but not least, I want to give my great gratitude to my family members, especially my parents. Without your support, I couldn't come to the Netherlands to fulfill my doctoral study (感谢爸妈和家人对我一贯以来的支持, 没有你们的支持和帮助, 我就不能完成博士研究以及追寻自己的梦想).

Curriculum Vitae

Born on 23rd March 1984 in Guangdong, China.

Graduated from high school (Chenghai High School, China) in 2003.

2003-2007: **BSc**, Faculty of Agriculture, Hainan University, Hainan, China, specializing in Biotechnology. Research project: Molecular cloning and analysis of mt16SrRNA in coconut leaf beetle (*Brontispa longissima*). Supervised by Prof. dr. Jun-sheng Huang (Chinese Academy of Tropical Agricultural Science).

2007-2010: **MSc**, Faculty of Life Science, Lanzhou University, Gansu, China, specializing in Animal Ecology. Research project: Life history evolution of black redstart in the northeastern Tibetan Plateau. Supervised by Prof. Naifa Liu (Lanzhou University).

2010-2014: **PhD** research project, Behavioral Biology group, Institute of Biology Leiden (IBL) Leiden University, Leiden, the Netherlands. Project title: Linguistic Birds – exploring the cognitive abilities in zebra finches by using artificial grammars. Supervised by Prof. dr. Carel J. ten Cate (Leiden University).

During this project, I supervised several BSc and MSc students and presented my work at national and international conferences (NVG: annual meeting of Netherlands Society for Behavioral Biology, IEC: International Ethological Conference, EVOLANG: Evolution of Language International Conference).

Publications

Scientific publications

Chen, J. N., Liu, N. F., Yan, C., & An, B. (2011). Plasticity in nest site selection of Black Redstart (*Phoenicurus ochruros*): a response to human disturbance. *Journal of Ornithology*, 152(3), 603-608.

Chen, J., van Rossum, D., & ten Cate, C. (2014). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition* (online, DOI: 10.1007/s10071-014-0786-4).

Chen, J., & ten Cate, C. (in press). Zebra finches can use positional and transitional cues to distinguish vocal element strings. *Behavioural Processes*

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(2), 165-175.

Manuscripts for publication

Chen, J., Jansen, N, & ten Cate, C. (under review). Zebra finch are able to learn prefixation and suffixation patterns.

Chen, J. & ten Cate, C. (in preparation). Bridging the gap: learning of non-adjacent dependencies in a songbird.