

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

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.

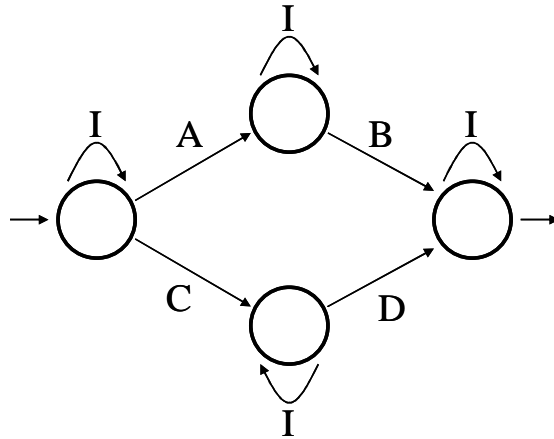


Fig. 1 Diagram of an artificial grammar that generates $I^n A I^{n2} B I^{n3}$ and $I^n C I^{n2} D I^{n3}$ ($n \geq 0$, 1-3 indicate that n is not necessary the same) sequences that are used in Experiments 1 and 2. Transitions from one state to another (arrow from one circle to another) produce an element 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).

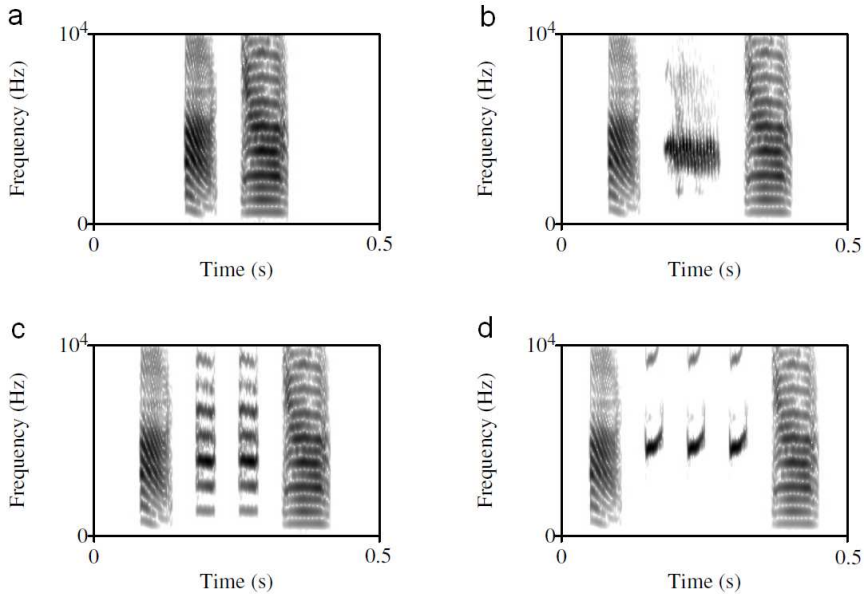


Fig. 2 Spectrograms of four examples of constructed song stimuli for bird 143 in Experiment 1. a) elements arranged in AB order. b) elements in AI₁B order. c) elements in AI₂I₂B order. d) elements in AI₃I₃I₃B order.

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 $G >$ 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).

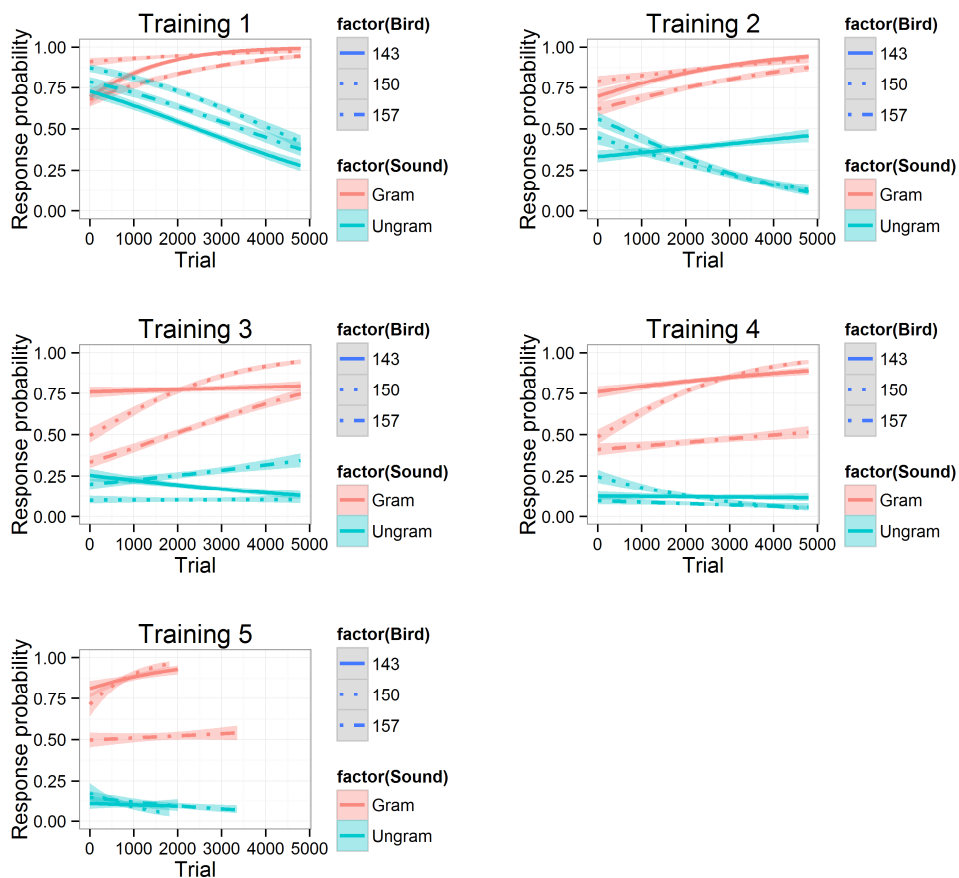


Fig. 3 Learning curves of the individual birds in Training 1-5, showing how the response to grammatical (indicated as Gram) and ungrammatical stimuli (indicated as Ungram) progressed over trials. Different line patterns and color indicate different individuals and sets of stimuli respectively. Broader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

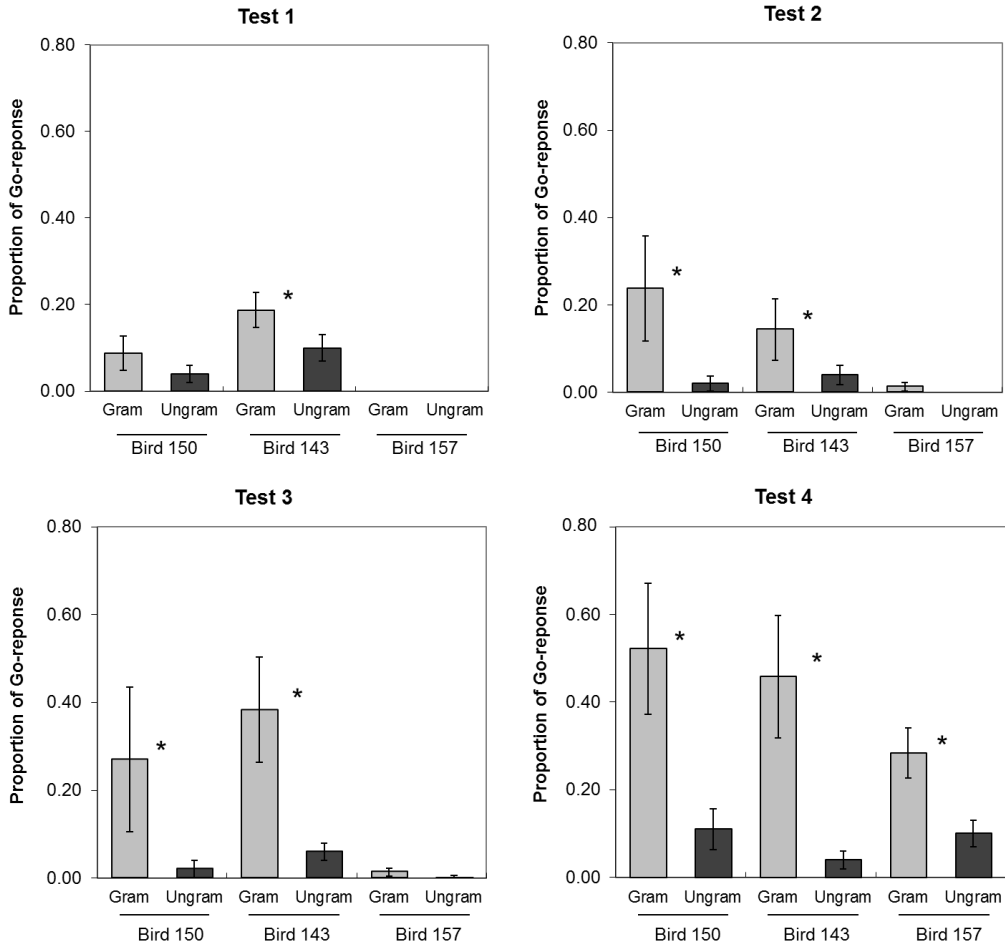


Fig. 4 Proportion of Go-responses of individual birds to grammatical (Gram) and ungrammatical stimuli (Ungram). A * indicates a significant difference between responses to grammatical and ungrammatical sets. In test 1, birds were exposed to test stimuli with one intervening item, such as AI₁B. In test 2, birds were exposed to test stimuli with 2 to 3 intervening items, like AI₁I₁B and AI₁I₁I₁B. In test 3 and 4, test stimuli contained novel intervening items (I₂ and I₃ 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).

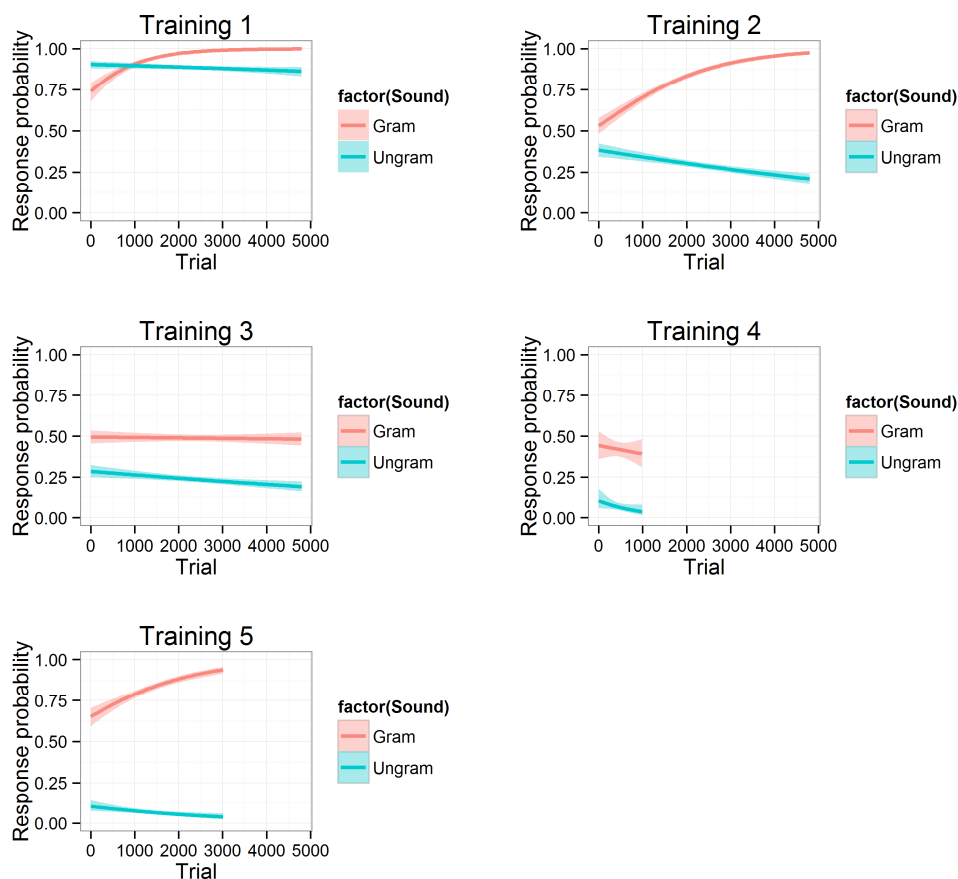


Fig. 5 Learning curves of bird 163 in Training 1-5, showing how response to grammatical (indicated as Gram) and ungrammatical stimuli (indicated as Ungram) progressed over trials. Different colors indicate different sets of stimuli. Broader color bands indicate the 95% confidence interval within one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

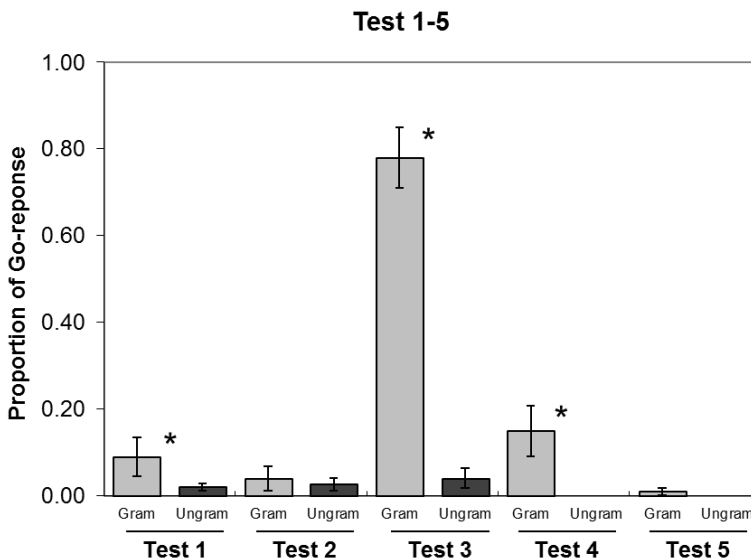


Fig. 6 Proportion of Go-responses of bird 163 to grammatical (Gram) and ungrammatical stimuli (Ungram). A * indicates a significant difference between responses to grammatical and ungrammatical sets. In test 1, the bird was exposed to test stimuli with one intervening item, such as A_1I_1B . In test 2, the bird was exposed to 5-element strings with dependencies occurring at arbitrary positions of the strings, like $A_1I_1B_1I_1I_1$ and $I_1A_1I_1B_1I_1$. 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/Nrn2931
- 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., Nespore, 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 CD	AD AC BD CB CA DB	1	AI ₁ B CI ₁ D	AI ₁ D CI ₁ B AI ₁ C BI ₁ D CI ₁ A DI ₁ B			
2	AI ₁ B CI ₁ D	AI ₁ D CI ₁ B AI ₁ C BI ₁ D CI ₁ A DI ₁ B	2	AI ₁ I 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 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	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	I ₁ AI ₁ DI ₁ CI ₁ BI ₁ I ₁ I ₁ CI ₁ I ₁ B I ₁ I ₁ DI ₁ B
3	AI ₁ I 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 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 ₂ B 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 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 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 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 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	4	AI ₃ B 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 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 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 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 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 I ₁ I ₁ D	3	AI ₁ I ₁ BI ₁	AI ₁ I ₁ DI ₁
	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	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		I ₁ AI ₁ I ₁ B CI ₁ I ₁ I ₁ D I ₁ CI ₁ DI 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	AI ₂ I ₂ I ₂ D	4	AI ₂ I ₂ I ₂ B	AI ₂ I ₂ I ₂ D
	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	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		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	I ₂ AI ₂ D 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	AI ₃ I ₃ I ₃ D	5	AI ₃ I ₃ I ₃ B	AI ₃ I ₃ I ₃ D
	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	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		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	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.

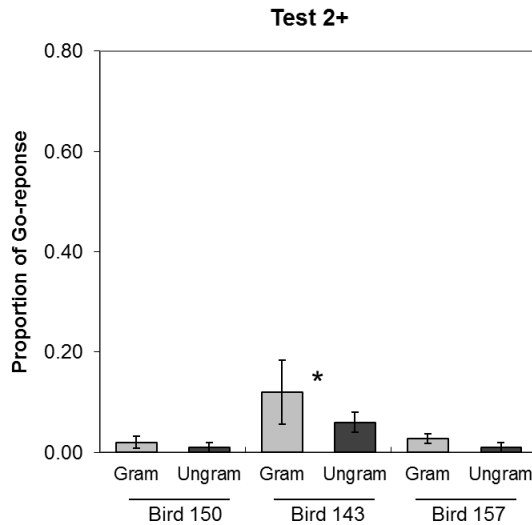
Training	Stimulus type & bird	Fitted Model
1	Ungram 143	Logit[Response]= 0.98518** - 0.000405** Trial
	Ungram 150	Logit[Response]= 1.927247** - 0.000468** Trial
	Ungram 157	Logit[Response]= 1.312305** - 0.00038** Trial
	Gram 143	Logit[Response]= 0.805128** + 0.000853** Trial
	Gram 150	Logit[Response]= 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.