Précis of *The music of language: Exploring grammar, prosody and rhythm perception in zebra finches and budgerigars.*

Michelle Spierings

1. Introduction

Humans are the only animal species known to communicate with the use of language, a highly-structured system which allows us to communicate limitless about intentions, experiences, thoughts and emotions. Although there are many non-human animals with elaborate communication systems, none of these systems have the complexity in rules and structure of human language. This does not mean, however, that non-human animals cannot learn different aspects of human language. Some of the cognitive requirements that enable language might actually be shared between species. These shared abilities did not evolve specifically for human language and might have been present in a common ancestor, preceding or enabling language evolution. Shared cognitive abilities can therefore inform us about the origin and evolutionary trajectory of language.

The faculty of language

The ability to perceive and produce language is not based one single cognitive trait. In fact, the faculty of language consists of a set of different cognitive mechanisms that work together[1-3]. Producing a variety of meaningful vocalizations, requires the cognitive ability to match the vocalizations produced by others. This vocal learning ability has so far only been found in humans and a few animal classes (cetaceans[4], pinnipeds[5-7], bats[8], elephants[9], hummingbirds[10], parrots[11] and songbirds[12]). Producing and perceiving these vocalizations in the correct way requires speech sound categorization and rule generalization. The first ability is essential to understand which sounds belong together to form a word and to know how different prosodic patterns change the meaning or intention of a word. The latter is required to abstract the structure by which sounds are combined into words and words are put together to form sentences. This includes the perception of grammatical rules.

The separation of the mechanisms underlying language in clear cognitive abilities allows us to study which abilities are uniquely human and which are shared with other species.

Comparative studies

When studying the evolutionary underpinnings of a human trait, it might seem most logical to study our closest relatives, the great apes. However, as mentioned earlier, one of the key components of language learning is vocal learning, an ability that has not been found in any non-human primate. The vocal learning ability also implies a strong focus from the learners on the vocal production of the teachers. This means that the auditory perception of vocal learners is well-developed and they are well equipped to perceive complex sounds and structures.

Of the vocal learning species, songbirds and parrots are two clades with the most complex vocal patterns. These vocalizations are learned from generation to generation and are used to recognize conspecifics and distinguish individuals[13]. Some songbirds have dialects depending on geographical location, which they can discriminate easily[14]. Moreover, the complex vocalizations are used in mate choice as they can reveal the quality of different individuals[15]. This shows that vocal learning birds can perceive and use variable, complex sounds, a useful starting point to study the cognitive underpinnings of language.

The most studied songbird species is the zebra finch. Both males and females produce social calls and the males sing one stereotyped song throughout their lives[16,17]. This song is a composition of elements from his fathers' song, combined with parts of songs from other males[18]. Zebra finches are closed ended learners, meaning that they learn their vocalizations when they are young, and are unable to learn new vocalizations later in life. Parrots, on the other hand, are open ended vocal learners, which means that they learn new vocalizations throughout their lives[19]. In budgerigars, one of the smaller parrot species, both males and females produce complex warble songs[20]. Budgerigars have been studied for their perception of human speech sounds and have similar sensitivities to phoneme boundaries as humans[21,22]. These two species, zebra finches and budgerigars, are the subjects of the experiments presented in this thesis.

This dissertation

This dissertation focusses on the question which cognitive abilities underlying language are shared between humans and non-human animals. The work is part of a collaborative project between developmental linguists, computational linguists and behavioural biologists. Together we compare the cognitive mechanisms related to language in human infants, adults and two bird species. Below, I will describe my part of this project, the study of the cognitive abilities in zebra finches and budgerigars, related to language perception. The work of this thesis will be described in the two general themes: speech sound perception and abstract structure generalization.

2. Speech sound perception in zebra finches

The most basic units of any language are the speech sounds. These speech sounds are stringed together to form words and sentences. Listeners need to be able to discriminate different speech sounds and to determine which sounds belong together as a word. Furthermore, the same speech sound can be produced with an increased pitch, duration or amplitude. This can be an indication of the in the structure of a word (in English, for example, words often start with a higher pitch syllable[23]), and hence enable the detection of word boundaries, or it can change the emotional content of a word or sentence. The following three experiments explored whether zebra finches have the abilities to perceive these aspects of lexical speech.

2.1 String segmentation

(Spierings, de Weger & ten Cate, Animal Cognition, 2015)

In natural language production, pauses between words are not as systematic as in written language. This makes it a complex task to determine where one word ends and the next one begins. In order to do this, infants use both prosodic features (pauses and word stress) as well as the co-occurrence of the syllables in a string[24-26].

In this experiment, we tested how zebra finches segment strings of zebra finch song elements. Just like infants, the birds could make use of the co-occurrence of the elements or (in a second group) also of the pauses between sets of elements (figure 1). They were trained to discriminate between two strings in a go/no-go task. These strings consisted of the same elements, but the way these elements co-occurred created different words in both strings. Next, they were asked to categorize words and mixtures of words from these training strings.

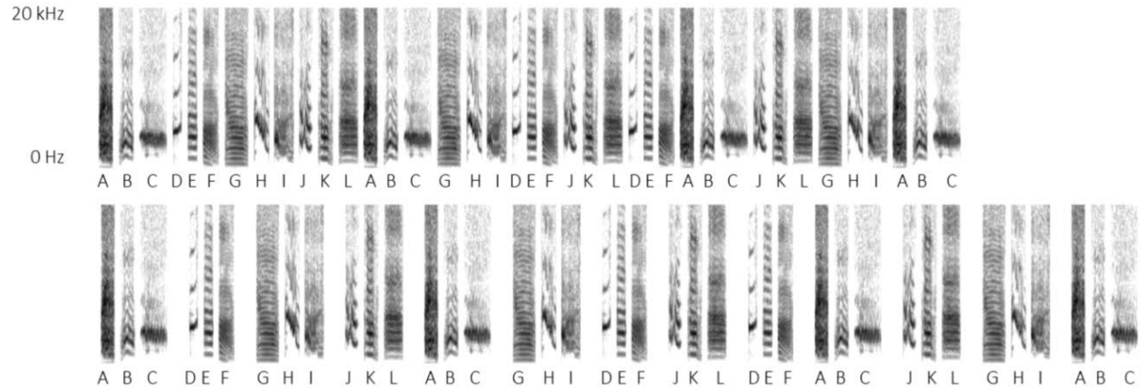

Fig. 1 Example of a string with and a string without pauses. Adjacent elements were separated by a 20 ms pause. In the Pause-condition a 80 ms pause separated the triplets.

The zebra finches that were trained with pauses between words were able to recognize the separate words of the training strings. The birds that could use co-occurrence, but not pauses, were unable to do so.

This shows that, unlike infants, zebra finches do not seem to make use of element co-occurrence to determine word boundaries. However, they do make use of the prosodic information (pauses), which is also a strategy used by humans. The next step is to determine if zebra finches are also sensitive to prosody in other language-related tasks. In the next two experiments, we study how zebra finches are perceiving and using prosodic information.

2.2 Perceptual grouping

(Spierings, Hubert & ten Cate, Animal Cognition, 2017)

Humans have a universal tendency to group strings of alternating tones in duplets with either initial (trochaic) or final (iambic) stress. For example, when hearing a string of tones alternating in pitch (high and low tones), humans perceive this as a concatenation of trochaic duplets[27]. When the tones are alternating in duration (long and short tones), humans group the tones as iambs or trochees, depending on their native language[27].

The only animal species thus far tested for perceptual grouping is the rat. These animals group strings alternating in pitch similar to humans, but do not show any perceptual grouping for duration alternations[28,29]. In our experiment, we studied whether zebra finches also have the tendency to perceive separate sounds as concatenated units. This sheds light on the question whether this is a mostly human ability, and the rats are a unique exception, or whether it is a more generally shared cognitive concept.

The zebra finches were trained to give a different behavioural response to quadruplets with a high-low or low-high pattern (group 1) or a long-short or short-long pattern (group 2). When they successfully made this discrimination, they were presented with long, ambiguous strings of alternating tones. Their responses to these long strings show whether they grouped them as concatenations of iambs or trochees.

Only the zebra finches that were trained on pitch alternations discriminated between the iambic and trochaic structures of the training. Interestingly, these individuals grouped the ambiguous strings as trochees, like humans. The birds that received the duration alternations did not manage to discriminate between iambs and trochees. This shows that not all sound modifications might result in perceptual grouping. Although songbirds are known to be sensitive to the general prosody of human speech[30], it remains a question to which features they are paying attention.

2.3 Prosody perception

(Spierings & ten Cate, Proc. R. Soc. B., 2014)

Prosody in human speech is paralinguistic information by which speakers can convey their emotions or alter the meaning of words and sentences. Prosody is created by an adjustment of the frequency, duration or amplitude of syllables. For example, the last word of a question usually has a higher frequency and a longer duration than the last word of an exclamation. This information is already used by infants to recognize their mother tongue[31] and to determine word boundaries[25]. All throughout our lives, prosody conveys important linguistic information.

Studies on the perception of human prosody by non-human animals show that they can discriminate language based on the prosodic patterns (rats[32], tamarins[33]). Java sparrows can even discriminate between the stress patterns of the same Japanese sentence spoken with admiration or suspicion[34]. This shows that there is certainly reason to assume that a sensitivity for prosodic changes might be shared more broadly in the animal kingdom. The question remains, however, how animals make these discriminations and which prosodic features they use.

In this study, we trained zebra finches to discriminate between two prosodic patterns and tested their sensitivity to different prosodic features. Zebra finches produce songs that are naturally varying in acoustic features. For example, they modify the amplitude of their songs when the receiver is further away[35] or change the duration of song syllables depending on the context[36]. Female zebra finches respond to these changes, indicating that they are sensitive to prosodic modulations.

The zebra finches were trained in a go/no-go setup to discriminate between xyxy and xxyy structures consisting of speech syllables, with prosodic stress on either the first or the last syllable (Xyxy versus xxyY). This stress was created by increases in the frequency, amplitude and duration of those syllables (figure 2). When the birds learned the discrimination, they were exposed to five different test categories to determine what they had learned. In test 1, the structural pattern and prosodic pattern were switched (Xxyy versus xyxY). In test 2, the prosodic information was reduced to only an increased frequency, duration or amplitude. Test 3 had quadruplets with two prosodic features increased on the first syllable and the third feature increased on the last syllable. In test 4 there was no prosodic pattern present, only the structural pattern. In the final test the syllables were changed to novel ones, to test whether the birds could generalize the prosody to new “words”.

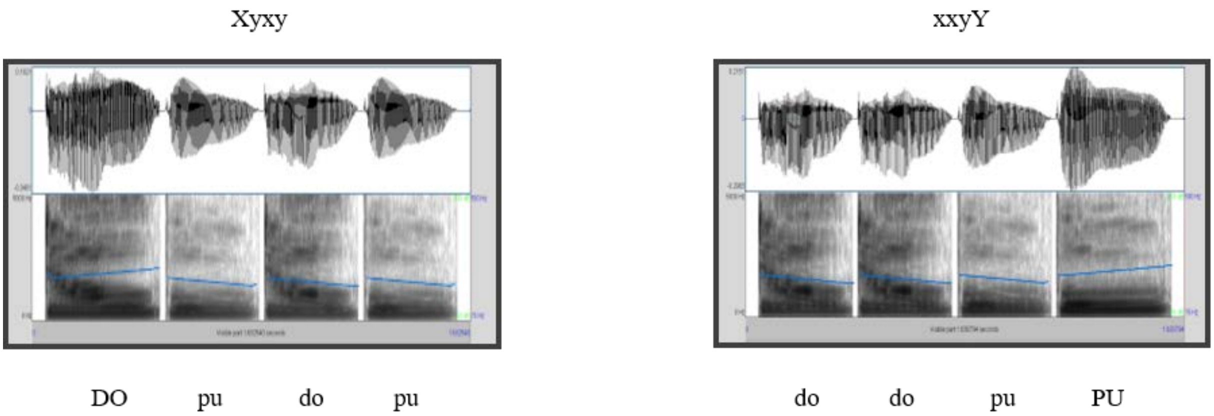

Fig. 2 Example of two training stimuli. ‘DO pu do pu’ follows the Xyxy structure and acts as a go-stimulus. The ‘do do pu PU’ stimulus follows the xxyY structure and is one of the no-go stimuli. In the sonogram, the change in frequency, amplitude and duration of the first or last syllable is visible.

Results show that zebra finches are fast in learning to discriminate between the stimuli and remain discriminating based on the prosody when the structure changes. This discrimination remains also when there is only one prosodic feature present, in which case they are most attentive to the increase in frequency. They can also generalize the prosodic pattern to novel syllables. In the absence of prosodic information, the zebra finches could still discriminate the quadruplets based on the structural pattern, but at a significantly lower rate.

This shows that the prosody of human speech is also salient to non-human animals. A sensitivity to prosody might therefore be one of the cognitive mechanisms that did not evolve specifically for human language.

3. Structure generalization in zebra finches and budgerigars

Crucial for language perception is the ability to understand the structure of sentences. This makes the difference between questions and statements for example. The cognitive capacity to learn abstract structures, enables us to generalize these structures to novel sounds. This cognitive ability has been difficult to show in non-human animals. The next two experiments are focussed on whether zebra finches and budgerigars can perceive, abstract and generalize structures in sound.

3.1 Rhythm perception

(Spierings & ten Cate, Frontiers in Neuroscience., 2016
ten Cate, Spierings, Hubert & Honing, Frontiers in Psychology, 2016)

Every song has a particular rhythm to it, but languages also have certain rhythms. Perceiving this rhythmicity requires the listener to form an abstract representation of the regularity of the occurrence of sounds. This abstraction should then allow for faster or slower varieties of the same rhythm. Humans are able to detect different rhythms, move their bodies in synchrony with the beat and recognize languages by their rhythmic features[37-39]. Especially infants have a strong sensitivity to rhythmic patterns in strings of sounds. They can easily discriminate between two rhythmic structures and generalize this to novel sounds[40]. Even newborns are already sensitive to regular beats and show an increased brain response to an omitted beat[41].

The study of rhythm perception in non-human animals has mainly focussed on rhythmic entrainment –the ability to move ones body simultaneously with a beat–. A sulphur-crested cockatoo was the first animal to show this. This bird entrained to a series of different pop-songs and stayed in synchrony with the beat when the songs were faster or slower[5,6]. Ever since, there have been several other examples of animals entraining to beats[42] resulting in the hypothesis that this ability might be linked to the vocal learning ability[43]. However, two species that are not known to be vocal learners, a Californian sea lion and a bonobo, also showed some rhythmic entrainment[44,45]. This opens up the question how widespread rhythm perception is, whether all vocal learners can do it and whether it might also be present in animals that are not able to control their body movements to show entrainment.

In this experiment, we tested both zebra finches and budgerigars on their rhythm perception abilities. Budgerigars are a representative of the parrot family, of which several species showed rhythmic entrainment. The zebra finches represent the songbird family, also vocal learners but with less clear entrainment abilities.

The birds were trained to discriminate between strings of tones that were either regular or irregular. After learning this discrimination, they were exposed to strings that were either faster or shorter, had longer or shorter tones, or longer or shorter pauses but were still regular and irregular. Both the zebra finches and the budgerigars learned to discriminate between the regular and the irregular strings. They generalized this regularity to some of the test strings, but were also paying attention to the exact duration of the tones and pauses.

This study shows that both zebra finches and budgerigars are able to generalize a rhythmic pattern as long as the tone or pause duration remains intact. There is no clear difference between the two species, indicating that parrots might not be particularly better at rhythmic perception then other vocal learners. Recently, it has been shown that zebra finches also produce rhythmic patterns and use the same brain regions when they are detecting and producing regular vocalizations[46-48]. We argue that this finding, together with our results reported above, make zebra finches an excellent model species for rhythm perception.

Combining our data with those of other bird species that we reviewed, we argue that there might not be such a clear difference between the rhythmic abilities of vocal learners and non-vocal learners. Rather, it appears that each species studied has a different level in how well they can generalize rhythmicity. It would therefore be a more plausible hypothesis that the rhythm perception ability (perhaps also the vocal learning ability) is not a matter of present or absent, but rather a graded scale.

3.2 Artificial grammar learning

(Spierings & ten Cate, PNAS, 2016)

In this last experiment, we studied one of the bedrocks of language, grammar learning. Listeners need to understand the grammatical rules of a particular language to understand the meaning of a sentence, as the order of words determines whether it is, for example, a question or a statement. The “Artificial Grammar Learning” (AGL) task was developed to systematically test grammar learning abilities in different experiments[49]. This paradigm has been used in many experiments with human adults and infants, showing that both exhibit impressive learning mechanisms in the visual and the auditory domain[50-53].

In a famous study, Marcus and colleagues[54] tested whether 8-month-old infants could learn simple XYX and XXY grammar rules (the X and the Y depict different speech syllables). The infants were familiarized with strings, consisting of speech syllables arranged in triplets in one of the two grammars, for only two minutes. When the infants heard both grammars with novel speech sounds afterwards, they showed an increased interest for the grammar that they had not been exposed to. This indicates that they quickly learned the abstract pattern of the grammars and were able to generalize this to novel sounds.

To examine the uniqueness of this ability, we studied whether zebra finches and budgerigars could learn to differentiate grammatical structures. The animals were trained to discriminate between 5 instances of an XYX grammar and 5 instances of an XXY grammar. These sounds consisted of zebra finch song elements and (in an additional zebra finch experiment) of artificial sounds. They were then tested on the same two grammars, now either consisting of novel sounds or of a combination of the training sounds. This allowed us to determine whether the animals were learning the individual sound strings of the triplets, or whether they were actually able to generalize the grammar to novel configurations and novel sounds.

The zebra finches learned to discriminate the XYX and XXY triplets during training, but in the test they were unable to generalize the pattern to novel sounds. Their responses to the novel configurations of training sounds showed that they were actually responding to the position of each sound item within the triplet. Although their ability to memorize the positions of the sounds is rather impressive, the tendency to use positional information is not uncommon in zebra finches. In earlier studies zebra finches used positional information when discriminating string of song elements[55]. To exclude the possibility that there was an influence of conspecific vocalizations on their learning strategy, the experiment was run again with artificial sounds. The results matched the previous experiment, showing that zebra finches indeed use positional information and not the general structure.

The budgerigars, on the other hand, did discriminate the test items based on their underlying structure. When given configurations of known sounds, the birds responded as they would to the training triplets. At first, they did not respond to the novel sounds and thus did not show any generalization of the grammar. In an additional test the budgerigars were only presented with novel items. This increased their response rate and showed that they were able to generalize the abstract pattern to grammars with novel sounds, more similarly to the infants.

This experiment is the first unambiguous evidence that the cognitive ability to learn and generalize abstract grammatical structures is not specific to humans or human language. Interestingly, similar experiments with infants show that they do not always generalize XYX and XXY strings based on their pattern[54], as the budgerigars do. When possible, infants also tend to use the positional information of the single items that is presented[56]. This is more similar to the responses of the zebra finches in our task. The behaviour of the two species thus reflects two different strategies, both of which are present in human infants. The ability to use positional information in a discrimination task is not unique in the animal kingdom[55, 57-59]. The ability to generalize abstract patterns, however, remains a rarity in the animal kingdom, perhaps suggesting a case of parallel evolution in humans and, at least, one non-human species.

4. Conclusion

Collectively, the work conducted in this thesis shows that there are multiple cognitive components of the language faculty that are not language specific, nor are they specific to humans. Zebra finches are capable of segmenting sound strings into “words” by paying attention to the pauses between the sounds[60]. Furthermore, they are sensitive to the prosodic features of human speech. They perceptually group tones alternating in pitch[61] and can discriminate, abstract and generalize the natural prosodic patterns of human speech[30]. Interestingly, similar abilities have been found in other species. Rats, for example, also perceptually group alternating tones and have a sensitivity for prosody[28,32] and some primate species can discriminate some prosodic features of speech sounds[33,38]. This shows that the cognitive mechanisms underlying our human abilities to perceive, discriminate and generalize prosodic speech features is shared throughout the animal kingdom. Possibly, our non-linguistic ancestors might already have been able to learn and categorize prosodic features. The evolution of language could have built upon this ability when language was shaped.

The cognitive ability to learn and generalize abstract structures, on the other hand, is more rare in animals. When both zebra finches and budgerigars were trained with rhythmic structures, their generalization ability was rather limited[62]. Furthermore, only the budgerigars could abstract and generalize grammatical structures[63]. There are hardly any convincing results of other animals showing similar abilities[64], indicating that this is not an ability that is widespread or shared between many species. The fact that both budgerigars and humans can learn grammars, might rather be a case of parallel evolution of this cognitive trait. This suggestion might, however, also be due to the limited amount of work that has been done on studying these abilities in non-human animals thus far[65]. However, the current state of the field suggests that the cognitive mechanism for abstraction is likely to have evolved later than the abilities related to prosody and speech perception.

Comparative studies, like the ones described in this thesis, will continue to clarify which cognitive mechanisms underpinning language are shared between species. The differences between species inform us whether certain other abilities, like vocal-learning, social structure, vocal production and recognition capacity, or cognitive abilities might be of importance for language perception. Moreover, these types of studies can also bridge the gap between music and language perception. The concept of language is merely human, and both language and music are a concatenation of sound items with a pattern of modifications. Studying how animals perceive these patterns and whether they are sensitive to particular musical or linguistic patterns, provides the basis to find the roots of music and language and their possible interaction. With more studies on closely and distantly related species, we can better pinpoint the cognitive mechanisms that might be underlying the evolution of music and language.

Taken together, this work provides a strong interdisciplinary foundation to develop systematic studies to further explore the cognitive abilities of animals that may have served as basis for the evolution of human language. These findings will together give us a better insight in the evolutionary process of language.

References

1. Jackendoff, R. (1997). The architecture of the language faculty. *MIT Press*.
   Johnson, E.K. & Seidl, A.H. (2009) At 11 months, prosody still outranks statistics. *Dev. Sci*., 12, 131 – 141.

2. Hauser, M.D., Chomsky, N. & Fitch, W.T. (2002). The Faculty of Language: What is it, who has it, and how did it evolve? *Science*, 298, 1569.

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

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

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

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

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

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

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

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

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

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

13. Lambrechts, M.M. & Dhondt, A.A. (1995). Individual voice recognition in birds. *Curr Ornithol.*, 12, 115-139.

14. Brooks, R.J. & Falls, J.B. (1975). Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbours and strangers. *Canadian Journal of Zoology*, 53, 879-888.

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

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

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

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

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

20. Farabaugh, S.M., Brown, E.D. & Dooling, R.J. (1992). Analysis of warble song of the budgerigar (*Melopsittacus undulates*). *The international journal of animal sound and its recording,* 4, 111-130.

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

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

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

24. Nazzi, T., Nelson, D.G.K., Jusczyk, P.W. & Jusczyk, A.M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. *Infancy*, 1(1), 123-147.

25. Thiessen, E.D. & Saffran, J.R. (2003). When cues collide: use of stress and statistical cue to word boundaries by 7-to 9-month old infants. *Dev. Psychol.*, 39(4), 706-716.

26. Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.

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

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

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

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

31. Johnson, E.K. & Seidl, A.H. (2009). At 11 months, prosody still outranks statistics. *Dev. Sci.*, 12, 131-141.

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

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

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

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

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

37. Merker, B.H., Madison, G.S. & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45, 4-17.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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, 669–680.

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

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

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

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

57. Scarf, D. & Colombo, M. (2010). Representation of serial order in pigeons (*Colomba livia*). *J. Exp. Psychol. Anim. Behav. Process.*, 36(4), 423-429.

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

59. Drucker, C.B. & Brannon, E.M. (2014). Rhesus monkeys (*Macaca mulatta*) map number onto space. *Cognition*, 132(1), 57-67.

60. Spierings, M., de Weger, A. & ten Cate, C. (2015). Pauses enhance chunk recognition in song element strings by zebra finches. *Animal Cognition*, 18, 867–874.

61. Spierings, M., Hubert, J., & ten Cate, C. (2017). Selective auditory grouping by zebra finches: Testing the Iambic-Trochaic law. *Animal Cognition*, 20(4), 665-675.

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

63. Spierings, M.J. & ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proc. Natl. Acad. Sci. USA*, 113(27), E3977–E3984.

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

65. ten Cate, C. (2016). Assessing the uniqueness of language: Animal grammatical abilities take center stage. *Psychon Bull Rev.*, 24(1), 91-96.